Two-Way Training Design for Discriminatory Channel Estimation in Wireless MIMO Systems

Prepared by Chao-Wei Huang

Directed by Prof. Yao-Win Peter Hong

In Partial Fulfillment of the Requirements

for the Degree of
Master of Science

Institute of Communications Engineering

National Tsing Hua University

Hsinchu, Taiwan 30013, R.O.C.

E-mail: cwhuang@erdos.ee.nthu.edu.tw

July, 2011

Abstract

This work examines the use of two-way training in multiple-input multiple-output (MIMO) wireless systems to discriminate the channel estimation performances between a legitimate receiver (LR) and an unauthorized receiver (UR). This thesis extends upon the previously proposed discriminatory channel estimation (DCE) scheme that allows only the transmitter to send training signals. The goal of DCE is to minimize the channel estimation error at LR while requiring the channel estimation error at UR to remain beyond a certain level. If the training signal is sent only by the transmitter, the performance discrimination between LR and UR will be limited since the training signals help both receivers perform estimates of their downlink channels. In this work, we consider instead the two-way training methodology that allows both the transmitter and LR to send training signals. In this case, the training signal sent by LR helps the transmitter obtain knowledge of the transmitter-to-LR channel, but does not help UR estimate its downlink channel (i.e., the transmitter-to-UR channel). With transmitter knowledge of the estimated transmitter-to-LR channel, artificial noise (AN) can then be embedded in the null space of the transmitter-to-LR channel to disrupt UR's channel estimation without severely degrading the channel estimation at LR. Based on these ideas, two-way DCE training schemes are developed for both reciprocal and non-reciprocal channels. The optimal power allocation between training and AN signals is devised under both average and individual power constraints. Numerical results are provided to demonstrate the efficacy of the proposed two-way DCE training schemes.

Contents

List of Figures

Chapter 1

Introduction

Secrecy in wireless communications has been an important problem over the years due to the broadcast nature of the wireless medium. In the past, these issues have mostly been addressed using cryptography in the application layer. However, recent studies on information-theoretic secrecy provide an alternative to achieve these tasks through coding and modulation in the physical layer. In the context of physical layer secrecy, one is often interested in deriving the so-called secrecy capacity, which is the rate achievable with vanishing error probability at the legitimate receiver (LR) and vanishing equivocation rate at the unauthorized receiver (UR). In particular, secrecy capacity has been derived for single-input single-output (SISO) systems in [1] and for multiple-input multiple-output (MIMO) systems in [2]. The results show that secrecy capacity can generally be increased by enlarging the difference between the effective channel qualities of LR and UR. While most works on physical layer secrecy focus on optimal coding schemes to achieve secrecy capacity in the data transmission phase, our goal is to exploit signal processing methods to enlarge the differences between the quality of the two channels. In particular, this is done from a channel estimation aspect, following the so-called *discriminatory channel estimation* (DCE) methodology proposed previously in [3].

Specifically, DCE is a training strategy that utilizes artificial noise (AN) to disrupt UR's reception while sending training signals to LR. In this case, AN must be placed in the null

space of the transmitter-to-LR channel to minimize its effect on LR. However, this requires transmitter knowledge of the channel, which is typically obtained through feedback from LR. In the original DCE scheme, only the transmitter is allowed to send training signals. In this case, increasing the training power helps improve the channel estimate at LR and allows for more effective use of AN at the transmitter. However, this also helps UR obtain a better channel estimate and, thus, the amount of power used for training must be limited. To improve the channel estimate at LR while confining the performance of UR to a certain level, multiple stages of feedback and retraining must be employed. In this case, training and AN power is increased as the transmitter knowledge of the channel improves through multiple stages. Yet, the training overhead and complexity required to optimize training over multiple stages limits its application in practice.

The main contribution of this thesis is to propose new and efficient DCE schemes using the two-way training methodology. Here, training signals will also be sent by LR and transmitter knowledge of the channel will be obtained by performing channel estimation at the transmitter. Notice that the original DCE scheme assumes that channel feedback with infinite resolution is provided from LR, which is not achievable in practice. When the channel is reciprocal, *e.g.*, in time-division multiplexing (TDD) systems, the channel state information (CSI) can be obtained at the transmitter by sending pilot signals from the receiver. Two-way training schemes have been studied for conventional point-to-point links in [4, 5, 6] to obtain the CSI at both the receiver and the transmitter without the use of feedback. In this work, we adopt the concept of two-way training into the design to increase the efficiency of the DCE scheme. In reciprocal channels, the proposed two-way DCE scheme uses reverse training to provide CSI at the transmitter and forward training with AN to achieve different channel estimation performances at LR and UR. When the channel is non-reciprocal, *e.g.*, in frequency-division multiplexing (FDD) systems, the downlink and uplink channels between

the transmitter and LR would not be identical. In this case, an additional training phase is needed, where the transmitter first broadcasts a randomly generated signal to LR, which then echoes the signal back to the transmitter. The echoed signal contains information of both the downlink and uplink channels and can be combined with the reverse training signal to estimate the desired transmitter-to-LR channel ($i.e.$, the downlink channel). Compared to the multi-stage feedback-and-retraining DCE scheme in [3], the proposed two-way training scheme drastically decreases the overall training overhead and design complexity.

To optimize the performance of the proposed two-way training scheme, we derive the optimal power allocation between the training and AN by solving an optimization problem that aims to minimize the channel estimation error at the LR subject to a lower limit constraint on the channel estimation error at the UR. In the reciprocal case, the analytical result shows that the problem of finding the optimal training and AN powers reduces to a one-variable problem which can be solved by a simple line search. However, in the non-reciprocal case, the power allocation problem is not easily solved since the estimation error expression is much more complex. Therefore, we instead resort to an approximate solution by using the monomial approximation and condensation method [14] in the field of geometric programming (GP). Numerical results show that the proposed DCE design can effectively discriminate the channel estimation and the data detection performances at the LR and UR.

The remainder of the thesis is organized as follows. In Chapter 2, we first introduce the wireless MIMO system model considered in this work and provide a general description of the DCE scheme. For the case with reciprocal channels, the training strategy is described in Chapter 3 and the optimal power allocation is derived in Chapter 4. Similarly, for the case with nonreciprocal channels, the training strategy is described in Chapter 5 and the optimal power allocation is derived in 6. Numerical results are provided in Chapter 7 and, finally, a conclusion is given in Chapter 8.

Chapter 2

System Model

Consider a wireless MIMO system that consists of a transmitter, a legitimate receiver (LR), and an unauthorized receiver (UR), as shown in Fig. 2.1. We assume that the transmitter, LR, and UR are equipped with N_t, N_L and N_U antennas, respectively. The channels of LR and UR remain constant during one transmission block, which consists of a training phase and a data transmission phase. The goal is to prevent the UR to extract information from its received signal. Instead of focusing the data transmission, we propose to achieve this task from a channel estimation perspective and devise two-way training schemes following the DCE methodology that enables LR to perform an accurate estimate of the channel while disrupting the channel estimation performance at UR.

Let the downlink channel from the transmitter to LR be denoted by $\mathbf{H}_d \in \mathbb{C}^{N_t \times N_L}$ and the uplink channel from LR to the transmitter be denoted by $\mathbf{H}_u \in \mathbb{C}^{N_L \times N_t}$. In the following chapters, we consider separately two different channel models, i.e., the reciprocal channel model and the non-reciprocal channel model. In both cases, the proposed two-way training scheme for DCE can be generally divided into two steps as described below.

Step I: The aim of Step I is to allow the transmitter to obtain an estimate of the downlink channel. Different from [3], where a noiseless feedback channel is required, we

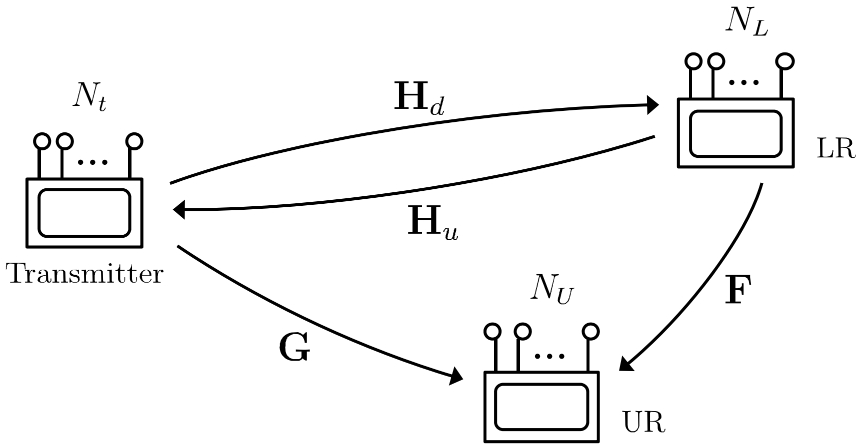


Figure 2.1: Diagram of a wireless MIMO system consisting of a transmitter, a legitimate receiver (LR) and an unauthorized receiver (UR).

allow the transmitter to estimate the downlink channel itself through the exchange of training signals between the transmitter and UR. This channel knowledge will be used in designing the forward training signal in order to discriminate the channel estimation performances between LR and UR. Different training strategies are required under different channel models to achieve the downlink channel estimation at the transmitter. Detailed descriptions for the reciprocal case and non-reciprocal case are given in Chapters 3 and 5, respectively.

Step II: After obtaining the downlink channel estimate in Step I, the transmitter next sends a training signal along with AN to degrade the channel estimation performance at UR. Specifically, by assuming that $N_t > N_L$, the forward training signal is given by

$$\mathbf{X}_t = \sqrt{\frac{\mathcal{P}_F \tau_F}{N_t}} \mathbf{C}_t + \mathbf{A}\mathbf{N}_{\widehat{H}_d}^H, \tag{2.1}$$

where $\mathbf{C}_t \in \mathbb{C}^{\tau_F \times N_t}$ is the pilot matrix with $\text{Tr}(\mathbf{C}_t^H \mathbf{C}_t) = N_t$, $\mathcal{P}_F$ is the forward training power, and τ_F training length. For ease of notation, we define $\mathcal{E}_F \triangleq \mathcal{P}_F \tau_F$ as the forward training energy. $\mathbf{A} \in \mathbb{C}^{\tau_F \times (N_t - N_L)}$ is the AN matrix of which each entry is $i.i.d.$ Gaussian with zero mean and variance σ_a^2 and $\mathbf{N}_{\widehat{H}_d} \in \mathbb{C}^{N_t \times (N_t - N_L)}$ is a matrix whose column vectors form an orthonormal basis for the left null space of $\widehat{\mathbf{H}}_d$, that is, $\mathbf{N}_{\widehat{H}_d}^H \widehat{\mathbf{H}}_d = \mathbf{0}_{(N_t - N_L) \times N_L}$ (i.e.,

the $(N_t - N_L)$ by N_L zero matrix) and $\mathbf{N}_{\widehat{H}_d}^H \mathbf{N}_{\widehat{H}_d} = \mathbf{I}_{N_t - N_L}$. Notice from (2.1) that AN is superimposed on the training signal and placed in the left null space of $\widehat{\mathbf{H}}_d$ to minimize its interference on LR. The received signals of the LR and UR are respectively given by

$$\mathbf{Y}_L = \sqrt{\frac{\mathcal{E}_F}{N_t}} \mathbf{C}_t \mathbf{H}_d + \mathbf{A}\mathbf{N}_{\widehat{H}_d}^H \mathbf{H}_d + \mathbf{W} \tag{2.2}$$

$$\mathbf{Y}_U = \sqrt{\frac{\mathcal{E}_F}{N_t}} \mathbf{C}_t \mathbf{G} + \mathbf{A}\mathbf{N}_{\widehat{H}_d}^H \mathbf{G} + \mathbf{V} \tag{2.3}$$

where $\mathbf{G} \in \mathbb{C}^{N_t \times N_U}$ is the channel matrix from the transmitter to UR, and $\mathbf{W} \in \mathbb{C}^{\tau_F \times N_L}$ and $\mathbf{V} \in \mathbb{C}^{\tau_F \times N_U}$ are the additive white noise matrices at LR and UR, respectively. Each entry of $\mathbf{G}$ is assumed to be $i.i.d.$ distributed with zero mean and variance equal to σ_G^2. Elements of both $\mathbf{W}$ and $\mathbf{V}$ are assumed to be $i.i.d.$ random variables with zero mean and variances respectively equal to σ_w^2 and σ_v^2.

In the following chapters, we describe the training strategies and examine the corresponding channel estimation performances at the transmitter and the receivers during each stage of the process. The optimal power allocation between training and AN signals are derived to achieve discrimination between the channel estimation performances at LR and UR.

Chapter 3

Two-Way Training Strategy for Reciprocal Channels

In this chapter, we consider the case where the channel between the transmitter and the LR is reciprocal, which means that the downlink and uplink channels are symmetric. In this case, we define the downlink channel matrix as $\mathbf{H}_d \triangleq \mathbf{H}$ and the uplink channel matrix as $\mathbf{H}_u \triangleq \mathbf{H}^T$. With channel reciprocity, the transmitter can obtain an estimate of the downlink channel by taking the transpose of the channel matrix obtained through reverse training, i.e., training based on signals sent from LR to the transmitter. In this case, DCE is effectively achieved using only two stages, i.e., the reverse and the forward training phases. The operations are detailed below.

Reverse Training: In the reverse training stage, LR first sends a training signal, denoted by $\mathbf{X}_L \in \mathbb{C}^{\tau_R \times N_L}$, to enable channel estimation at the transmitter. Specifically, the reverse training signal $\mathbf{X}_L$ is given by

$$\mathbf{X}_L = \sqrt{\frac{\mathcal{P}_R \tau_R}{N_L}} \mathbf{C}_L, \tag{3.1}$$

where the pilot matrix $\mathbf{C}_L$ satisfies $\mathbf{C}_L^H \mathbf{C}_L = \mathbf{I}_{N_L}$ (the N_L by N_L identity matrix), and $\mathcal{P}_R$ and τ_R represent the transmission power and training interval, respectively. For ease of use later, we define the reverse training as $\mathcal{E}_R \triangleq \mathcal{P}_R \tau_R$. The received signal at the transmitter is

given by

$$\mathbf{Y}_t = \mathbf{X}_L \mathbf{H}^T + \widetilde{\mathbf{W}}, \tag{3.2}$$

where each element of $\mathbf{H}$ is assumed to be independent and identically distributed (i.i.d.) random variable with zero mean and variance equal to σ_H^2, and $\widetilde{\mathbf{W}} \in \mathbb{C}^{\tau_R \times N_t}$ is the additive white noise matrix with each element having zero mean and variance $\sigma_{\tilde{w}}^2$. By the help of reverse training, the channel estimate of $\mathbf{H}$, denoted by $\widehat{\mathbf{H}}$, can be obtained at the transmitter

Forward Training: In the forward training stage, the transmitter superimposes AN on top of the training signal to degrade the channel estimation performance at UR. With knowledge of the estimated downlink channel, i.e., $\widehat{\mathbf{H}}$, AN can be placed in the left null space of $\widehat{\mathbf{H}}$ to minimize its interference on the LR. The forward training signal is given in (2.1) where the pilot matrix $\mathbf{C}_t$ satisfies $\mathbf{C}_t^H \mathbf{C}_t = \mathbf{I}_{N_t}$. And the received signals of LR and UR are respectively given in (2.2) and (2.3). Note that the notation $\widehat{\mathbf{H}}_d$ in (2.1) and (2.2) is replaced by $\widehat{\mathbf{H}}$.

In the next chapter, we analyze the channel estimation performances of transmitter, LR and UR by assuming that all of them employ linear minimum mean square error (LMMSE) criterion for channel estimation [7]. We then propose to judiciously allocate the training powers and the AN power in reverse and forward training, aiming at discriminating between the channel estimation performances of the LR and the UR.

Chapter 4

Optimal Power Allocation for DCE in Reciprocal Channels

4.1 Channel Estimation Performance at Transmitter

Due to channel reciprocity, the reverse training signals sent by LR allow the transmitter to obtain an estimate of the downlink channel by taking the transpose of its estimate of the uplink channel. By employing the LMMSE estimator, the estimate of the channel matrix $\mathbf{H}$ can be written as

$$\begin{aligned}\widehat{\mathbf{H}} &= (\sigma_H^2 \mathbf{X}_L^H (\sigma_H^2 \mathbf{X}_L \mathbf{X}_L^H + \sigma_{\tilde{w}}^2 \mathbf{I}_{\tau_R})^{-1} \mathbf{Y}_t)^T \\ &\triangleq \mathbf{H} + \Delta\mathbf{H} \end{aligned} \tag{4.1}$$

where $\Delta\mathbf{H} \in \mathbb{C}^{N_t \times N_L}$ stands for the estimation error matrix. The covariance matrix of $\Delta\mathbf{H}$ can be shown to be [7]

$$\mathrm{E}\{\Delta\mathbf{H}(\Delta\mathbf{H})^H\} = N_L \left(\frac{1}{\sigma_H^2} + \frac{\mathcal{E}_R}{N_L \sigma_{\tilde{w}}^2} \right)^{-1} \mathbf{I}_{N_t}. \tag{4.2}$$

4.2 Channel Estimation Performance at LR and UR

To analyze the channel estimation performance of LR, let us write (2.2) as

$$\mathbf{Y}_L = \sqrt{\frac{\mathcal{E}_F}{N_t}}\mathbf{C}_t\mathbf{H} - \mathbf{A}\mathbf{N}_{\widehat{H}}^H\Delta\mathbf{H} + \mathbf{W} \triangleq \bar{\mathbf{C}}\mathbf{H} + \bar{\mathbf{W}}, \tag{4.3}$$

where $\bar{\mathbf{C}} \triangleq \sqrt{\frac{\mathcal{E}_F}{N_t}}\mathbf{C}_t$, $\bar{\mathbf{W}} \triangleq -\mathbf{A}\mathbf{N}_{\widehat{H}}^H\Delta\mathbf{H} + \mathbf{W}$ and the first equality is due to $\mathbf{N}_{\widehat{H}}^H\widehat{\mathbf{H}} = \mathbf{0}$. Denote the channel estimate at LR by $\widehat{\mathbf{H}}_L$. The normalized mean squared error (NMSE) of $\widehat{\mathbf{H}}_L$ under LMMSE criterion can be shown to be [7]

$$\begin{aligned}
\mathrm{NMSE}_L &\triangleq \frac{\mathrm{Tr}\left(\mathrm{E}\{(\mathbf{H}-\widehat{\mathbf{H}}_L)(\mathbf{H}-\widehat{\mathbf{H}}_L)^H\}\right)}{N_tN_L} \\
&= \frac{\mathrm{Tr}\left(\left(\mathbf{R}_H^{-1} + \bar{\mathbf{C}}^H\mathbf{R}_{\bar{W}}^{-1}\bar{\mathbf{C}}\right)^{-1}\right)}{N_tN_L},
\end{aligned} \tag{4.4}$$

where $\mathrm{Tr}(\cdot)$ denotes the trace of a matrix, $\mathbf{R}_H = N_L\sigma_H^2\mathbf{I}_{N_t}$ and $\mathbf{R}_{\bar{W}} = \mathrm{E}\{\mathbf{W}\bar{\mathbf{W}}^H\}$ is the covariance matrix of $\bar{\mathbf{W}}$. According to the independence between $\mathbf{A}$ and $\mathbf{W}$, the fact of $\mathbf{N}_{\widehat{H}}^H\mathbf{N}_{\widehat{H}} = \mathbf{I}_{N_t-N_L}$ and (4.2), it can be shown that

$$\begin{aligned}
\mathbf{R}_{\bar{W}} &= \left(\mathrm{E}\{\|\mathbf{N}_{\widehat{H}}^H\Delta\mathbf{H}\|^2\}\sigma_a^2 + N_L\sigma_w^2\right)\mathbf{I}_{\tau_F} \\
&= N_L\left[(N_t-N_L)\cdot\left(\frac{1}{\sigma_H^2} + \frac{\mathcal{E}_R}{N_L\sigma_{\tilde{w}}^2}\right)^{-1}\sigma_a^2 + \sigma_w^2\right]\mathbf{I}_{\tau_F}.
\end{aligned} \tag{4.5}$$

Substituting (4.5) into (4.4) yields

$$\begin{aligned}
\mathrm{NMSE}_L &= \frac{\mathrm{tr}\left(\frac{1}{N_L\sigma_H^2}\mathbf{I}_{N_t} + \frac{\mathcal{E}_F}{N_tN_L}\frac{\mathbf{C}_t^H\mathbf{C}_t}{(N_t-N_L)\left(\frac{1}{\sigma_H^2}+\frac{\mathcal{E}_R}{N_L\sigma_{\tilde{w}}^2}\right)^{-1}\sigma_a^2+\sigma_w^2}\right)^{-1}}{N_tN_L} \\
&= \left(\frac{1}{\sigma_H^2} + \frac{\mathcal{E}_F/N_t}{(N_t-N_L)\left(\frac{1}{\sigma_H^2}+\frac{\mathcal{E}_R}{N_L\sigma_{\tilde{w}}^2}\right)^{-1}\sigma_a^2+\sigma_w^2}\right)^{-1}.
\end{aligned} \tag{4.6}$$

The NMSE performance of the UR can be analyzed in a similar way. Specifically, one can show that the NMSE of estimating $\mathbf{G}$ at the UR is given by

$$\mathrm{NMSE}_U = \left(\frac{1}{\sigma_G^2} + \frac{\mathcal{E}_F/N_t}{(N_t-N_L)\sigma_a^2\sigma_G^2+\sigma_v^2}\right)^{-1}. \tag{4.7}$$

4.3 Optimal Power Allocation between Training and AN Signals

Observing from (4.6) and (4.7), the added AN in forward training can affect both the LR and UR's channel estimation performances. To optimize LR's channel estimation performance while preventing the UR from obtaining an accurate estimate of $\mathbf{G}$, we propose to jointly optimize the reverse training energy $\mathcal{E}_R$, the forward training energy $\mathcal{E}_F$ and AN power σ_a^2 by considering the following power allocation problem

$$
\begin{aligned}
\min_{\mathcal{E}_R,\mathcal{E}_F,\sigma_a^2\geq 0} \quad & \mathrm{NMSE}_L \\
\text{s.t.} \quad & \mathrm{NMSE}_U \geq \gamma, \\
& \mathcal{E}_R + \mathcal{E}_R + (N_t - N_L)\sigma_a^2\tau_F \leq P_{ave}(\tau_R + \tau_F), \\
& \mathcal{E}_R \leq \bar{P}_L\tau_R, \\
& \mathcal{E}_F + (N_t - N_L)\sigma_a^2\tau_F \leq \bar{P}_t\tau_F,
\end{aligned}
\tag{4.8}
$$

where we aim to minimize the LR's NMSE subject to a preset lower limit γ on the UR's NMSE, under an average power constraint P_{ave}. Note that the LR and the transmitter also have their own peak power constraints, i.e., $\bar{P}_L$ and $\bar{P}_t$.

Remark: In the DCE scheme, it is desirable to keep the forward training length as small as possible, *i.e.*, equal to the number of transmit antennas. Observing from (4.6), (4.7) and the problem (4.8) and assuming the average energy constraint and the individual energy constraint on the transmitter and the LR are all fixed, as the forward training length increases, it needs more AN energy to meet the same lower limit value and thus less energy can be allocated to the training signal. Different from the receiver's noise of which the energy can be freely accumulated over time, it takes the system's resource to maintain the AN's power.

To make all constraints effective, we shall focus on the interesting case where

$$\max\{\bar{P}_L\tau_R,\ \bar{P}_t\tau_F\} \leq P_{ave}(\tau_R+\tau_F) \leq \bar{P}_L\tau_R+\bar{P}_t\tau_F. \tag{4.9}$$

Note that for the case where $P_{ave}(\tau_R+\tau_F) > \bar{P}_L\tau_R+\bar{P}_t\tau_F$, the average power constraint becomes redundant and hence, the transmitter and the LR simply transmit with its maximum power. When $P_{ave}(\tau_R+\tau_F) < \bar{P}_L\tau_R$ and/or $P_{ave}(\tau_R+\tau_F) < \bar{P}_t\tau_F$, one or both individual power constraints become redundant. The solution for this case can be easily obtained by following the derivations for the case of (4.9) [1].

On the other hand, it should be noted that the preset value γ should satisfy [3]

$$\left(\frac{1}{\sigma_G^2}+\frac{\min\{\bar{P}_t\tau_F, P_{ave}(\tau_R+\tau_F)\}}{N_t\sigma_v^2}\right)^{-1} \leq \gamma \leq \sigma_G^2, \tag{4.10}$$

since the left-hand-side term is the minimum achievable NMSE of UR (when the transmitter does not use AN, i.e., $\sigma_a^2=0$), and the right-hand-side term stands for the worst NMSE performance of UR, respectively. For ease of latter use, let us define

$$\tilde{\gamma} \triangleq \left(\frac{1}{\gamma}-\frac{1}{\sigma_G^2}\right)N_t\sigma_v^2 \geq 0. \tag{4.11}$$

Then the condition in (4.10) reduces to

$$0 \leq \tilde{\gamma} \leq \bar{P}_t\tau_F. \tag{4.12}$$

The power allocation problem in (4.8) is a nonconvex optimization problem involving three variables $(\mathcal{E}_R, \mathcal{E}_F, \sigma_a^2)$. However, it actually can be solved very efficiently. We show in Appendix 9.1 the following proposition for problem (4.8):

Proposition 1. *Consider the power allocation problem in* (4.8) *with both* (4.9) *and* (4.12) *satisfied. If*

$$\mu \triangleq N_L\left(\frac{\sigma_v^2\sigma_{\tilde{w}}^2}{\sigma_G^2\sigma_w^2}-\frac{\sigma_{\tilde{w}}^2}{\sigma_H^2}\right) > \min\{\bar{P}_L\tau_R,\ P_{ave}(\tau_R+\tau_F)-\tilde{\gamma}\}$$

[1] The proposition to be given for the case of (4.9) also describes the solution for the case of $P_{ave}(\tau_R+\tau_F) < \bar{P}_L\tau_R$ and/or $P_{ave}(\tau_R+\tau_F) < \bar{P}_t\tau_F$, by changing the condition in (4.12) to $0 \leq \tilde{\gamma} \leq \min\{\bar{P}_t\tau_F,\ P_{ave}(\tau_R+\tau_F)\}$ and setting the redundant individual power constraint(s) to infinity.

then the optimal $(\mathcal{E}_R, \mathcal{E}_F, \sigma_a^2)$ *of* (4.8) *is given by* $\mathcal{E}_R* = 0$, $\mathcal{E}_F^* = \tilde{\gamma}$ *and* $(\sigma_a^2)^* = 0$ *(i.e., no need of reverse training and no need of AN in forward training). On the other hand, if* $\mu \leq \min\{\bar{P}_L \tau_R,\ P_{ave}(\tau_R + \tau_F) - \tilde{\gamma}\}$, *the optimal* $\mathcal{E}_R$ *of* (4.8) *can be obtained by solving the following one-dimensional problem*

$$
\begin{aligned}
\mathcal{E}_R^\star = \arg \max_{\mathcal{E}_R \geq 0} \quad & \frac{(N_L \sigma_{\tilde{w}}^2 + \sigma_H^2 \mathcal{E}_R) b(\mathcal{E}_R)}{N_L \sigma_{\tilde{w}}^2 + \sigma_H^2 \cdot \mathcal{E}_R + N_L \sigma_H^2 \frac{\sigma_{\tilde{w}}^2}{\sigma_w^2} \cdot c(\mathcal{E}_R)} \\
\text{s.t} \quad & \max\{0, \mu, P_{ave}(\tau_R + \tau_F) - \bar{P}_t \tau_F\} \leq \mathcal{E}_R \leq \\
& \min\{\bar{P}_L \tau_R,\ P_{ave}(\tau_R + \tau_F) - \tilde{\gamma}\},
\end{aligned}
\tag{4.13}
$$

where

$$
\alpha(\mathcal{E}_R) = \frac{\bar{P}_a ve(\tau_R + \tau_F) - \tilde{\gamma} - \mathcal{E}_R}{\tau_F + \sigma_G^2 \tilde{\gamma} / \sigma_v^2}, \tag{4.14}
$$

and

$$
\mathcal{E}_F(\mathcal{E}_R) = \tilde{\gamma} \left(\frac{\sigma_G^2}{\sigma_v^2} \cdot \alpha(\mathcal{E}_R) + 1 \right). \tag{4.15}
$$

The optimal $\mathcal{E}_F$ *and* σ_a^2 *are given by* $\mathcal{E}_F^\star = \mathcal{E}_F(\mathcal{E}_R^\star)$ *and* $(\sigma_a^2)^\star = \frac{\alpha(\mathcal{E}_R^\star)}{(N_t - N_L)}$.

Proposition 1 implies that the solutions of problem (4.8) can be efficiently obtained by simple line search over a finite interval, when the condition in (4.13) is fulfilled; otherwise, one can obtain a simple closed-form solution of $\mathcal{E}_R^* = 0$, $\mathcal{E}_F^* = \tilde{\gamma}$ and $(\sigma_a^2)^* = 0$.

Chapter 5

Two-Way Training Strategy for Non-reciprocal Channels

In this chapter, we consider the case of non-reciprocal channels, where the downlink and uplink channel matrices are asymmetric. In this case, the downlink channel cannot be directly inferred from the uplink channel. Therefore, an additional training stage using an echoed signal (from transmitter to LR and back to the transmitter) is needed in order to obtain an estimate of the combined downlink and uplink channel. This additional stage is referred to as round-trip training. The proposed two-way training method for DCE in non-reciprocal case is detailed below.

Round-trip Training: In round-trip training, the transmitter first broadcasts a random signal then the LR will echo its received signal back. By the round-trip procedure, the echoed signal obtained at the transmitter contains a combined term of uplink channel and downlink channel. Then with the help of the following reverse training, the transmitter can obtain the downlink channel estimate. Specifically, the random signal sent by the transmitter is given by

$$\mathbf{X}_{t0} = \sqrt{\frac{\mathcal{P}_0 \tau_0}{N_t}} \mathbf{C}_{t0}, \tag{5.1}$$

where $\mathbf{C}_{t0} \in \mathbb{C}^{\tau_0 \times N_t}$ is the pilot matrix satisfying $\mathrm{Tr}(\mathbf{C}_{t0}^H \mathbf{C}_{t0}) = N_t$, and $\mathcal{P}_0$ and τ_0 represent

the training power and training length, respectively. For ease of use later, we define the round-trip training energy as $\mathcal{E}_0 \triangleq \mathcal{P}_0\tau_0$. The received signal at the LR is given by

$$\mathbf{Y}_{L0} = \mathbf{X}_{t0}\mathbf{H}_d + \mathbf{W}_0, \tag{5.2}$$

where each element of $\mathbf{H}_d$ is assumed to be $i.i.d.$ complex Gaussian random variable with zero mean and variance equal to $\sigma^2_{H_d}$ and $\mathbf{W}_0 \in \mathbb{C}^{\tau_0 \times N_L}$ is the additive white Gaussian noise (AWGN) matrix with each entry having zero mean and variance σ^2_w. Then the LR amplifies and forwards its received signal back to the transmitter. The echoed signal at the transmitter is given by

$$\begin{aligned}\mathbf{Y}_{t1} &= \alpha\mathbf{Y}_{L0}\mathbf{H}_u + \widetilde{\mathbf{W}}_1 \qquad (5.3)\\ &= \alpha\mathbf{X}_{t0}\mathbf{H}_d\mathbf{H}_u + \alpha\mathbf{W}_0\mathbf{H}_u + \widetilde{\mathbf{W}}_1\end{aligned}$$

where each element of $\mathbf{H}_u$ is assumed to be $i.i.d.$ complex Gaussian random variables with zero mean and variance $\sigma^2_{H_u}$, $\widetilde{\mathbf{W}}_1 \in \mathbb{C}^{\tau_0 \times N_t}$ is the AWGN matrix at the transmitter with the power of each entry equal to $\sigma^2_{\tilde{w}}$. The amplifying gain at the LR is given by

$$\begin{aligned}\alpha &= \sqrt{\frac{\mathcal{P}_1\tau_0}{\mathcal{P}_0\tau_0 N_L\sigma^2_{H_d} + \tau_0 N_L \sigma^2_w}}\\ &= \sqrt{\frac{\mathcal{E}_1}{\mathcal{E}_0 N_L\sigma^2_{H_d} + \tau_0 N_L \sigma^2_w}} \qquad (5.4)\end{aligned}$$

where $\mathcal{P}_1$ is the transmission power and $\mathcal{E}_1 \triangleq \mathcal{P}_1\tau_0$ is the energy of the transmitted symbol. Since the random signal $\mathbf{X}_{t0}$ is available at the transmitter, it is able to obtain the downlink channel estimate with a given uplink channel. We will see how a reverse training helps the transmitter to extract the knowledge of downlink channel $\mathbf{H}_d$. Note that the random signal $\mathbf{X}_{t0}$ is unknown to both LR and UR, therefore the UR can not benefit from the round-trip training.

Reverse Training: In reverse training, the LR sends a training signal $\mathbf{X}_{L2} \in \mathbb{C}^{\tau_2 \times N_L}$ to enable the uplink channel estimation at the transmitter. Specifically, the reverse training

signal is given by

$$\mathbf{X}_{L2} = \sqrt{\frac{\mathcal{P}_2\tau_2}{N_L}}\mathbf{C}_{L2}, \tag{5.5}$$

where $\mathbf{C}_{L2}$ is the pilot matrix which satisfies $\mathrm{Tr}(\mathbf{C}_{L2}^H\mathbf{C}_{L2}) = N_L$, and $\mathcal{P}_2$ and τ_2 is the transmission power and training interval of the LR. For simplicity, We define the reverse training energy as $\mathcal{E}_2 \triangleq \mathcal{P}_2\tau_2$. The received signal at the transmitter is given by

$$\mathbf{Y}_{t2} = \mathbf{X}_{L2}\mathbf{H}_u + \widetilde{\mathbf{W}}_2 \tag{5.6}$$

where $\widetilde{\mathbf{W}}_2 \in \mathbb{C}^{\tau_2 \times N_t}$ is the additive white noise matrix with each entry having zero mean and variance $\sigma_{\tilde{w}}^2$. As the uplink channel estimate is given, the downlink channel estimate can be acquired from the echoed signal.

Forward Training: In forward training, the transmitter sends AN along with the training signal to discriminate the channel estimation performances between LR and UR. The specific description is stated in the Step II of Chapter 2. Note that we replace the subscript by 3 for notation consistency in this chapter, therefore the received signals at the LR and UR are replaced by

$$\mathbf{Y}_{L3} = \sqrt{\frac{\mathcal{E}_3}{N_t}}\mathbf{C}_{t3}\mathbf{H}_d + \mathbf{A}\mathbf{N}_{\widehat{H}_d}^H\mathbf{H}_d + \mathbf{W}_3 \tag{5.7}$$

$$\mathbf{Y}_{U3} = \sqrt{\frac{\mathcal{E}_3}{N_t}}\mathbf{C}_{t3}\mathbf{G} + \mathbf{A}\mathbf{N}_{\widehat{H}_d}^H\mathbf{G} + \mathbf{V}_3 \tag{5.8}$$

and the forward training length τ_3 is to substitute τ_F.

Due to the complicated nature of the two-way training, finding the optimal pilot structures may be a difficult task, which could also be different for different objective functions, *e.g.* channel estimation error, bit error rate or ergodic capacity, *etc.* The practical intuition in choosing the pilot structure is 1) to reduce the channel estimation error and 2) to reduce the transmission overhead. In conventional channel estimation, the orthogonal structure was usually found to be good. Note that it may not be the optimal choice for the system

we are considering. By utilizing the orthogonal training signal, the performance of channel estimation is now determined by the training energy of each phase and one can keep the training length minimum if the training energy can be designed to reduce the channel estimation error. Besides, it is preferred to keep the training length small in the secrecy channel estimation according to the remark stated in Chapter 4. Hence, in this work, we choose the minimum training length to be the number of transmit antenna, $i.e.$, $\tau_0 = \tau_3 = N_t$ and $\tau_2 = N_L$. And we assume the unitary pilot data are used, that is $\mathbf{C}_{t0}^H\mathbf{C}_{t0} = \mathbf{C}_{t0}\mathbf{C}_{t0}^H = \mathbf{I}_{N_t}$, $\mathbf{C}_{L2}^H\mathbf{C}_{L2} = \mathbf{C}_{L2}\mathbf{C}_{L2}^H = \mathbf{I}_{N_L}$ and $\mathbf{C}_{t3}^H\mathbf{C}_{t3} = \mathbf{C}_{t3}\mathbf{C}_{t3}^H = \mathbf{I}_{N_t}$.

Chapter 6

Optimal Power Allocation for DCE in Nonreciprocal Channels

In this chapter, we show how the transmitter can compute the downlink channel estimate from the training signals and analyze the channel estimation performance at both LR and UR. We assume that the transmitter, LR and UR all employ the linear minimum mean square error (LMMSE) criterion for channel estimation [7]. Then, we examine the optimal power allocation between training and AN signals in this case and propose an efficient solution for this problem.

6.1 Channel Estimation Performance at Transmitter

In this section , we show how the transmitter computes the downlink channel estimate from the reverse and round-trip training signals. Specifically, with the help of reverse training and by employing the LMMSE estimator, the estimate of the uplink channel $\mathbf{H}_u$ can first be computed as [7]

$$\widehat{\mathbf{H}}_u = \sigma_{H_u}^2 \mathbf{X}_{L2}^H (\sigma_{H_u}^2 \mathbf{X}_{L2} \mathbf{X}_{L2}^H + \sigma_{\tilde{w}}^2 \mathbf{I}_{N_t})^{-1} \mathbf{Y}_{t2} \triangleq \mathbf{H}_u + \Delta \mathbf{H}_u \tag{6.1}$$

where $\Delta\mathbf{H}_u \in \mathbb{C}^{N_L \times N_t}$ is the estimation error matrix with correlation matrix given by

$$\mathrm{E}\{(\Delta\mathbf{H}_u)^H \Delta\mathbf{H}_u\} = N_L \left(\frac{1}{\sigma_{H_u}^2} + \frac{\mathcal{E}_2}{N_L \sigma_{\tilde{w}}^2} \right)^{-1} \mathbf{I}_{N_t}. \tag{6.2}$$

With the uplink channel estimate $\widehat{\mathbf{H}}_u$ being available at the transmitter, we can rewrite the echoed signal (5.3) as

$$\begin{aligned} \mathbf{Y}_{t1} &= \alpha \mathbf{X}_{t0} \mathbf{H}_d (\widehat{\mathbf{H}}_u - \Delta\mathbf{H}_u) + \alpha \mathbf{W}_0 (\widehat{\mathbf{H}}_u - \Delta\mathbf{H}_u) + \widetilde{\mathbf{W}}_1 \qquad (6.3) \\ &= \alpha \mathbf{X}_{t0} \mathbf{H}_d \widehat{\mathbf{H}}_u + (-\alpha \mathbf{W}_0 \widehat{\mathbf{H}}_u - \alpha \mathbf{X}_{t0} \mathbf{H}_d \Delta\mathbf{H}_u + \alpha \mathbf{W}_0 \Delta\mathbf{H}_u + \widetilde{\mathbf{W}}_1). \end{aligned}$$

To employ the LMMSE criterion for the downlink channel estimation at the transmitter, it is easier to empress (6.3) in the vector form as

$$\mathbf{y}_{t1} = \alpha(\widehat{\mathbf{H}}_u^T \otimes \mathbf{X}_{t0})\mathbf{h}_d - \alpha(\Delta{\mathbf{H}_u}^T \otimes \mathbf{X}_{t0})\mathbf{h}_d + \alpha(\widehat{\mathbf{H}}_u^T \otimes \mathbf{I}_{N_t})\mathbf{w}_0 - \alpha(\Delta{\mathbf{H}_u}^T \otimes \mathbf{I}_{N_t})\mathbf{w}_0 + \tilde{\mathbf{w}}_1 \tag{6.4}$$

where the fact that $\mathrm{vec}(ABC) = (C^T \otimes A)\mathrm{vec}(B)$ is used, $\mathbf{y}_{t1} = \mathrm{vec}(\mathbf{Y}_{t1})$ is formed by stacking the columns of $\mathbf{Y}_{t1}$ and so do $\mathbf{h}_d = \mathrm{vec}(\mathbf{H}_d)$, $\mathbf{w}_0 = \mathrm{vec}(\mathbf{W}_0)$, and $\tilde{\mathbf{w}}_1 = \mathrm{vec}(\widetilde{\mathbf{W}}_1)$. As $\widehat{\mathbf{H}}_u$ is given at the transmitter, by the fact that $\widehat{\mathbf{H}}_u$ and $\Delta\mathbf{H}_u$ are uncorrelated due to the orthogonality principle [7], the premise of $\mathbf{C}_{t0}\mathbf{C}_{t0}^H = \mathbf{C}_{t0}^H\mathbf{C}_{t0} = \mathbf{I}_{N_t}$ and (6.2), the LMMSE estimate of downlink channel $\mathbf{h}_d$ and thus its matrix form are respectively given by

$$\begin{aligned} \hat{\mathbf{h}}_{d,t} &= \frac{1}{\alpha\sigma_w^2} \left(\frac{1}{\sigma_{H_d}^2} + \frac{\mathcal{E}_0}{N_t \sigma_w^2} \right)^{-1} \left(\widehat{\mathbf{H}}_u^* \left((\widehat{\mathbf{H}}_u^T \widehat{\mathbf{H}}_u^*) + \beta \mathbf{I}_{N_t} \right)^{-1} \otimes \mathbf{X}_{t0}^H \right) \mathbf{y}_{t1} && (6.5) \\ &\triangleq \mathbf{h}_d + \Delta\mathbf{h}_{d,t} && (6.6) \\ \widehat{\mathbf{H}}_{d,t} &= \frac{1}{\alpha\sigma_w^2} \left(\frac{1}{\sigma_{H_d}^2} + \frac{\mathcal{E}_0}{N_t \sigma_w^2} \right)^{-1} \mathbf{X}_{t0}^H \mathbf{Y}_{t1} \left((\widehat{\mathbf{H}}_u^H \widehat{\mathbf{H}}_u) + \beta \mathbf{I}_{N_t} \right)^{-1} \widehat{\mathbf{H}}_u^H && (6.7) \\ &\triangleq \mathbf{H}_d + \Delta\mathbf{H}_{d,t} && (6.8) \end{aligned}$$

where

$$\beta = N_L \left(\frac{1}{\sigma_{H_u}^2} + \frac{\mathcal{E}_2}{N_L \sigma_{\tilde{w}}^2} \right)^{-1} + \frac{\sigma_{\tilde{w}}^2}{\alpha^2 \sigma_{H_d}^2 \sigma_w^2} \left(\frac{1}{\sigma_{H_d}^2} + \frac{\mathcal{E}_0}{N_t \sigma_w^2} \right)^{-1} \tag{6.9}$$

and $\Delta \mathbf{h}_{d,t} \in \mathbb{C}^{N_t N_L \times 1}$ is the estimation error vector at the transmitter. The correlation matrix of $\Delta \mathbf{h}_{d,t}$ conditioned on a given $\widehat{\mathbf{H}}_u$ is given by

$$\mathrm{E}\{\Delta \mathbf{h}_{d,t}(\Delta \mathbf{h}_{d,t})^H|\widehat{\mathbf{H}}_u\} = \left[\sigma_{H_d}^2 \mathbf{I}_{N_L} - \sigma_{H_d}^2 \frac{\sigma_{H_d}^2 \mathcal{E}_0}{\sigma_{H_d}^2 \mathcal{E}_0 + N_t \sigma_w^2} \left(\left(\frac{1}{\beta} \widehat{\mathbf{H}}_u^* \widehat{\mathbf{H}}_u^T \right)^{-1} + \mathbf{I}_{N_L} \right)^{-1} \right] \otimes \mathbf{I}_{N_t} \tag{6.10}$$

Note that for differentiating from the downlink channel estimate of the LR, we denote the downlink channel estimate of the transmitter as $\widehat{\mathbf{H}}_{d,t}$. The matrix consisting of the basis of left null space of $\widehat{\mathbf{H}}_{d,t}$ is replaced as $\mathbf{N}_{\widehat{H}_{d,t}}$.

6.2 Channel Estimation Performance at LR and UR

In this section we analyze the channel estimation performance of the LR and UR. We first consider the channel estimation at the LR. Due to the fact that $\mathbf{N}_{\widehat{H}_d}^H \widehat{\mathbf{H}}_{d,t} = \mathbf{0}$ the received signal of LR (5.7) can be written as

$$\mathbf{Y}_{L3} = \bar{\mathbf{C}}_{t3} \mathbf{H}_d - \mathbf{A} \mathbf{N}_{\widehat{H}_{d,t}}^H \Delta \mathbf{H}_{d,t} + \mathbf{W}_3. \tag{6.11}$$

where $\bar{\mathbf{C}}_{t3} \triangleq \sqrt{\frac{\mathcal{E}_3}{N_t}} \mathbf{C}_{t3}$. To apply the LMMSE criterion for the downlink channel estimation of the LR, let us vectorize (6.11) as

$$\mathbf{y}_{L3} = \left(\mathbf{I}_{N_L} \otimes \bar{\mathbf{C}}_{t3}\right) \mathbf{h}_d - (\mathbf{I}_{N_L} \otimes \mathbf{A} \mathbf{N}_{\widehat{H}_d}^H) \Delta \mathbf{h}_{d,t} + \mathbf{w}_3 \tag{6.12}$$

where $\mathbf{h}_{d,t} = \mathrm{vec}(\mathbf{H}_d)$ and $\mathbf{w}_3 = \mathrm{vec}(\mathbf{W}_3)$. Then the channel estimate of $\mathbf{h}_d$ is given by

$$\hat{\mathbf{h}}_d = \mathbf{C}_{h_d y_{L3}} \mathbf{C}_{y_{L3} y_{L3}}^{-1} \mathbf{y}_{L3} \tag{6.13}$$

where

$$\mathbf{C}_{h_d y_{L3}} = \mathrm{E}\{\mathbf{h}_d \mathbf{y}_{L3}^H\} = \sigma_{H_d}^2 \left(\mathbf{I}_{N_L} \otimes \bar{\mathbf{C}}_{t3}\right) \tag{6.14}$$

is the covariance matrix between $\mathbf{h}_d$ and $\mathbf{y}_{L3}$ and

$$\mathbf{C}_{y_{L3}y_{L3}} = \mathrm{E}\{\mathbf{y}_{L3}\mathbf{y}_{L3}^H\} \tag{6.15}$$

$$= \sigma_{H_d}^2\left(\mathbf{I}_{N_L}\otimes\bar{\mathbf{C}}_{t3}\bar{\mathbf{C}}_{t3}^H\right) + \mathrm{E}\{(\mathbf{I}_{N_L}\otimes\mathbf{A}\mathbf{N}_{\widehat{H}_d}^H)\Delta\mathbf{h}_{d,t}\Delta\mathbf{h}_{d,t}^H(\mathbf{I}_{N_L}\otimes\mathbf{A}\mathbf{N}_{\widehat{H}_d}^H)^H\} + \sigma_w^2(\mathbf{I}_{N_L}\otimes\mathbf{I}_{N_t}) \tag{6.16}$$

is the covariance matrix of $\mathbf{y}_{L3}$. The expectation in (6.16) is taken over all the random variables including $\mathbf{A}$, $\Delta\mathbf{h}_{d,t}$ and $\widehat{\mathbf{H}}_{d,t}$ of which the last two are functions of the random matrix $\widehat{\mathbf{H}}_u$. With the law of iterated expectations, *i.e.*, $\mathrm{E}\{X\} = \mathrm{E}\{\mathrm{E}\{X|Y\}\}$, the second term of (6.16) can be written as

$$\mathrm{E}_{\widehat{\mathbf{H}}_u}\{\mathrm{E}_{\mathbf{A},\widehat{\mathbf{H}}_{d,t}}\{(\mathbf{I}_{N_L}\otimes\mathbf{A}\mathbf{N}_{\widehat{H}_{d,t}}^H)\mathrm{E}\{\Delta\mathbf{h}_{d,t}\Delta\mathbf{h}_{d,t}^H|\widehat{\mathbf{H}}_{d,t},\widehat{\mathbf{H}}_u\}(\mathbf{I}_{N_L}\otimes\mathbf{A}\mathbf{N}_{\widehat{H}_d}^H)^H|\widehat{\mathbf{H}}_u\}\} \tag{6.17}$$

where a fact that the random matrix $\mathbf{A}$ is independent of $\Delta\mathbf{h}_{d,t}$ is used. Since the term $\mathrm{E}\{\Delta\mathbf{h}_{d,t}\Delta\mathbf{h}_{d,t}^H|\widehat{\mathbf{H}}_{d,t},\widehat{\mathbf{H}}_u\}$ is not easy to tackle, we made an assumption that $\widehat{\mathbf{H}}_{d,t}$ is Gaussian distributed under a given $\widehat{\mathbf{H}}_u$. In this case, $\Delta\mathbf{H}_{d,t} = \mathbf{H}_d - \widehat{\mathbf{H}}_{d,t}$ is also Gaussian distributed and so does its vector form $\Delta\mathbf{h}_{d,t}$. We know that $\Delta\mathbf{H}_{d,t}$ is uncorrelated to $\widehat{\mathbf{H}}_{d,t}$ refering to the orthogonality principle, therefore $\Delta\mathbf{h}_{d,t}$ and $\widehat{\mathbf{H}}_{d,t}$ are independent due to our imposed Gaussian assumption. The equation in (6.17) is then given by

$$\mathrm{E}_{\widehat{\mathbf{H}}_u}\{\mathrm{E}_{\mathbf{A},\widehat{\mathbf{H}}_{d,t}}\{(\mathbf{I}_{N_L}\otimes\mathbf{A}\mathbf{N}_{\widehat{H}_{d,t}}^H)\mathrm{E}\{\Delta\mathbf{h}_{d,t}\Delta\mathbf{h}_{d,t}^H|\widehat{\mathbf{H}}_u\}(\mathbf{I}_{N_L}\otimes\mathbf{A}\mathbf{N}_{\widehat{H}_d}^H)^H|\widehat{\mathbf{H}}_u\}\} \tag{6.18}$$

Substituting (6.10) and the fact that $\mathbf{N}_{\widehat{H}_{d,t}}^H\mathbf{N}_{\widehat{H}_{d,t}} = \mathbf{I}_{N_t-N_L}$ into (6.18), we obtain

$$(N_t-N_L)\sigma_a^2\left[\sigma_{H_d}^2\mathbf{I}_{N_L} - \sigma_{H_d}^2\frac{\sigma_{H_d}^2\mathcal{E}_0}{\sigma_{H_d}^2\mathcal{E}_0+\sigma_w^2}\mathrm{E}\left\{\left(\left(\frac{1}{\beta}\widehat{\mathbf{H}}_u^*\widehat{\mathbf{H}}_u^T\right)^{-1}+\mathbf{I}_{N_L}\right)^{-1}\right\}\right]\otimes\mathbf{I}_{N_t} \tag{6.19}$$

The Hermitian term $\widehat{\mathbf{H}}_u\widehat{\mathbf{H}}_u^H$ can be factorized into

$$\widehat{\mathbf{H}}_u\widehat{\mathbf{H}}_u^H = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \tag{6.20}$$

where $\mathbf{U}\in\mathbb{C}^{N_L\times N_L}$ is the matrix whose columns are consisting of eigenvectors of $\widehat{\mathbf{H}}_u\widehat{\mathbf{H}}_u^H$ and $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1,\ldots,\lambda_{N_L})$ is the diagonal matrix with diagonal elements being nonzero and

unordered eigenvalues of $\widehat{\mathbf{H}}_u\widehat{\mathbf{H}}_u^H$. Since the elements of both the uplink channel $\mathbf{H}_u$ and the noise matrix $\widetilde{\mathbf{W}}_2$ are $i.i.d.$ Gaussian distributed and the reverse training $\mathbf{X}_{L2}$ is assumed to be semi-unitary, each entry of $\widehat{\mathbf{H}}_u$ is $i.i.d.$ Gaussian distributed. Referring to [10], we know that $\widehat{\mathbf{H}}_u\widehat{\mathbf{H}}_u^H$ has a Wishart distribution with N_t degrees of freedom and its mean is given by

$$\mathrm{E}\{\widehat{\mathbf{H}}_u\widehat{\mathbf{H}}_u^H\} = N_t \frac{\sigma_{H_u}^4 \mathcal{E}_2}{\sigma_{H_u}^2 \mathcal{E}_2 + N_L \sigma_{\tilde{w}}^2} \mathbf{I}_{N_L} \triangleq N_t \sigma^2 \mathbf{I}_{N_L}$$

where

$$\sigma^2 \triangleq \frac{\sigma_{H_u}^4 \mathcal{E}_2}{\sigma_{H_u}^2 \mathcal{E}_2 + N_L \sigma_{\tilde{w}}^2} \tag{6.21}$$

is the variance of each $i.i.d.$ random variable of $\widehat{\mathbf{H}}_u$. Both the density function of $\widehat{\mathbf{H}}_u\widehat{\mathbf{H}}_u^H$ [10] and the Jacobian of the eigenvalue value decomposition of $\widehat{\mathbf{H}}_u\widehat{\mathbf{H}}_u^H$ [11] can be divided into the product of functions of $\mathbf{\Lambda}$ and $\mathbf{U}$, thus we conclude that $\mathbf{\Lambda}$ and $\mathbf{U}$ are independent. With the independency and applying the law of iterated expectations, the equation in (6.19) becomes

$$(N_t - N_L)\sigma_a^2 \left[\sigma_{H_d}^2 \mathbf{I}_{N_L} - \sigma_{H_d}^2 \frac{\sigma_{H_d}^2 \mathcal{E}_0}{\sigma_{H_d}^2 \mathcal{E}_0 + N_t \sigma_w^2} \mathrm{E}_{\mathbf{U}}\left\{\mathbf{U} \cdot \mathrm{E}_{\mathbf{\Lambda}}\left\{\left(\beta \mathbf{\Lambda}^{-1} + \mathbf{I}_{N_L}\right)^{-1}\right\}\mathbf{U}^H\right\}\right] \otimes \mathbf{I}_{N_t} \tag{6.22}$$

$$= (N_t - N_L)\sigma_a^2 \left[\sigma_{H_d}^2 - \sigma_{H_d}^2 \frac{\sigma_{H_d}^2 \mathcal{E}_0}{\sigma_{H_d}^2 \mathcal{E}_0 + N_t \sigma_w^2} \mathrm{E}_{\lambda_1}\left\{\left(\frac{1}{\beta/\lambda_1 + 1}\right)\right\}\right] \mathbf{I}_{N_L} \otimes \mathbf{I}_{N_t} \tag{6.23}$$

where the equality holds since the eigenvalues of the Wishart distributed matrix $\widehat{\mathbf{H}}_u\widehat{\mathbf{H}}_u^H$ have identical distributions as any one of the unordered eigenvalues [12]. Replacing (6.23) in (6.16), we have an approximation of the covariance matrix of $\mathbf{y}_{L3}$ as

$$\mathbf{C}_{y_{L3}y_{L3}} \approx$$

$$\mathbf{I}_{N_L} \otimes \left\{\sigma_{H_d}^2 \bar{\mathbf{C}}_{t3} \bar{\mathbf{C}}_{t3}^H + \left[(N_t - N_L)\sigma_a^2 \left(\sigma_{H_d}^2 - \sigma_{H_d}^2 \frac{\sigma_{H_d}^2 \mathcal{E}_0}{\sigma_{H_d}^2 \mathcal{E}_0 + N_t \sigma_w^2} \mathrm{E}_{\lambda_1}\left\{\left(\frac{1}{\beta/\lambda_1 + 1}\right)\right\}\right) + \sigma_w^2\right] \mathbf{I}_{N_t}\right\} \tag{6.24}$$

The normalized mean squared error (NMSE) of $\widehat{\mathbf{H}}_d$ can be computed as

$$\mathrm{NMSE}_L = \frac{\mathrm{Tr}(\mathrm{E}\{\Delta\mathbf{h}_d\Delta\mathbf{h}_d^H\})}{N_tN_L} \tag{6.25}$$

$$= \frac{\mathrm{Tr}\left(\sigma_{H_d}^2\mathbf{I}_{N_LN_t} - \mathbf{C}_{h_dy_{L3}}\mathbf{C}_{y_{L3}y_{L3}}^{-1}\mathbf{C}_{h_dy_{L3}}^H\right)}{N_tN_L} \tag{6.26}$$

Substituting (6.14) and (6.24) into (6.26), we have an approximation for the NMSE of the LR as

$$\mathrm{NMSE}_L \approx \left(\frac{1}{\sigma_{H_d}^2} + \frac{\mathcal{E}_3}{N_t}\frac{1}{(N_t-N_L)\sigma_a^2\left(\sigma_{H_d}^2 - \sigma_{H_d}^2\frac{\sigma_{H_d}^2\mathcal{E}_0}{\sigma_{H_d}^2\mathcal{E}_0+N_t\sigma_w^2}\mathrm{E}_{\lambda_1}\left\{\left(\frac{1}{\beta/\lambda_1+1}\right)\right\}\right)+\sigma_w^2}\right)^{-1} \tag{6.27}$$

For $N_t \gg 1$, the distribution of the eigenvalues of $\widehat{\mathbf{H}}_u\widehat{\mathbf{H}}_u^H$ is asymptotically approximated to a Gaussian distribution [13], that is $\lambda_1 \overset{a.}{\sim} \mathcal{N}(N_t\sigma^2, N_t\sigma^4)$ where σ^2 is given in (6.21). However, the expectation term in (6.27) is intractable, we instead apply the Jensen's inequality and take its lower bound as an approximation. Hence, we have

$$\mathrm{NMSE}_L \approx \left(\frac{1}{\sigma_{H_d}^2} + \frac{\mathcal{E}_3}{N_t}\frac{1}{(N_t-N_L)\sigma_a^2\left(\sigma_{H_d}^2 - \sigma_{H_d}^2\frac{\sigma_{H_d}^2\mathcal{E}_0}{\sigma_{H_d}^2\mathcal{E}_0+N_t\sigma_w^2}\left(\frac{N_t\sigma}{\beta+N_t\sigma}\right)\right)+\sigma_w^2}\right)^{-1} \tag{6.28}$$

On the other hand, the NMSE performance of the UR is analyzed as follows. The received signal of UR (5.8) can be vectorized as

$$\mathbf{y}_{U3} = \mathrm{vec}(\mathbf{Y}_{U3}) = (\mathbf{I}_{N_U}\otimes\bar{\mathbf{C}}_{t3})\mathbf{g} + (\mathbf{I}_{N_U}\otimes\mathbf{A}\mathbf{N}_{\widehat{H}_d}^H)\mathbf{g} + \mathbf{v}_3 \tag{6.29}$$

where $\bar{\mathbf{C}}_{t3} = \sqrt{\frac{\mathcal{E}_3}{N_t}}\mathbf{C}_{t3}$, $\mathbf{g} = \mathrm{vec}(\mathbf{G})$ and $\mathbf{v}_3 = \mathrm{vec}(\mathbf{V}_3)$. The covariance matrix between $\mathbf{g}$ and $\mathbf{y}_{U3}$ and the covariance matrix of $\mathbf{y}_{U3}$ are respectively given by

$$\mathbf{C}_{g,y_{U3}} = \sigma_G^2(\mathbf{I}_{N_U}\otimes\bar{\mathbf{C}}_{t3}^H) \tag{6.30}$$

$$\mathbf{C}_{y_{U3}y_{U3}} = \mathbf{I}_{N_U}\otimes\left[\sigma_G^2\bar{\mathbf{C}}_{t3}\bar{\mathbf{C}}_{t3}^H + \left(\sigma_G^2(N_t-N_L)\sigma_a^2+\sigma_v^2\right)\mathbf{I}_{N_t}\right] \tag{6.31}$$

Hence, the NMSE of the UR is given by

$$
\begin{aligned}
\text{NMSE}_U &= \frac{\text{Tr}(\sigma_G^2 \mathbf{I}_{N_t N_U} - \mathbf{C}_{g,y_{U3}} \mathbf{C}_{y_{U3}y_{U3}}^{-1} \mathbf{C}_{g,y_{U3}}^H)}{N_t N_U} \\
&= \frac{\text{Tr}\left(\mathbf{I}_{N_U} \otimes \left\{ \sigma_G^2 \mathbf{I}_{N_t} - \sigma_G^2 \bar{\mathbf{C}}_{t3} \left[\sigma_G^2 \bar{\mathbf{C}}_{t3} \bar{\mathbf{C}}_{t3}^H + (\sigma_G^2 (N_t - N_L) \sigma_a^2 + \sigma_v^2) \mathbf{I}_{N_t} \right]^{-1} \sigma_G^2 \bar{\mathbf{C}}_{t3} \right\} \right)}{N_t N_U} \\
&= \left(\frac{1}{\sigma_G^2} + \frac{\mathcal{E}_3}{N_t} \frac{1}{\sigma_G^2 (N_t - N_L) \sigma_a^2 + \sigma_v^2} \right)^{-1} \qquad (6.32)
\end{aligned}
$$

6.3 Optimal Power Allocation between Training and AN Signals

With (6.28) and (6.32), we can jointly design the power values of $\{\mathcal{E}_0,\ \mathcal{E}_1,\ \mathcal{E}_2,\ \mathcal{E}_3,\ \sigma_a^2\}$ by considering the following power allocation problem

$$
\begin{aligned}
\min_{\mathcal{E}_0,\mathcal{E}_1,\mathcal{E}_2,\mathcal{E}_3,\sigma_a^2 \geq 0} \quad & \text{NMSE}_L && (6.33a) \\
\text{s.t.} \quad & \text{NMSE}_U \geq \gamma && (6.33b) \\
& \mathcal{E}_0 + \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 + (N_t - N_L)\sigma_a^2 N_t \leq P_{ave}(N_t + N_t + N_L + N_t) && (6.33c) \\
& \mathcal{E}_0 + \mathcal{E}_3 + (N_t - N_L)\sigma_a^2 N_t \leq \bar{P}_t (N_t + N_t) && (6.33d) \\
& \mathcal{E}_1 + \mathcal{E}_2 \leq \bar{P}_L (N_t + N_L). && (6.33e)
\end{aligned}
$$

Here, we aim to minimize the NMSE of LR subject to the constraint that the NMSE of UR remains above a preset lower limit γ. We also consider the average power constraint P_{ave} and two individual power constraints $\bar{P}_t$ and $\bar{P}_L$ at the transmitter and LR, respectively. However, the problem is not easily solvable. To obtain an efficient solution, we resort to the monomial approximation and the condensation method often adopted in the field of geometric programming (GP) [14]. Details are given in the Appendix 9.2.

Chapter 7

Numerical Results and Discussions

In this chapter, we present numerical results to demonstrate the effectiveness of the proposed DCE schemes. We consider the MIMO wireless system as described in Chapter 2 with $N_t = 4$, $N_L = 2$ and $N_U = 2$. The elements of the channel matrices $\mathbf{H}$ and $\mathbf{G}$ are $i.i.d.$ complex Gaussian distributed with zero mean and unit variance ($\sigma_H^2 = \sigma_G^2 = 1$). Each entry of additive white noise matrices $\widetilde{\mathbf{W}}$, $\mathbf{W}$ and $\mathbf{V}$ is also $i.i.d.$ complex Gaussian distributed with zero mean and unit variance, $i.e.$, $\sigma_{\tilde{w}}^2 = \sigma_w^2 = \sigma_v^2 = 1$. Moreover, the training lengths are set to be the antenna number of the terminal which transmits that training signal, $i.e.$, $\tau_R = N_L = 2$ and $\tau_F = N_t = 4$ for the reciprocal case[1] and $\tau_0 = \tau_3 = N_t = 4$ and $\tau_2 = N_L = 2$ for the non-reciprocal case. Note that the overall training time is larger than the sum of all training length due to the processing time at the transmitter. Besides, the individual power constraints of the transmitter and the LR are respectively assigned as $\bar{P}_t = 30$ dB and $\bar{P}_L = 20$ dB. We incorporate an NMSE lower bound for comparison. The lower bound for reciprocal and non-reciprocal case are respectively given by

$$\mathrm{NMSE}_{\mathrm{LB},rec} = \left(\frac{1}{\sigma_H^2} + \frac{\min\{\bar{P}_t N_t,\ P_{ave}(N_L + N_t)\}}{N_t \sigma_w^2}\right)^{-1} \tag{7.1}$$

[1]In pure channel estimation, it is preferred to keep the training length minimum in uncorrelated channel and white noise [8]. We show in Figure 7.1 that the training length is better to choose as smallest length $i.e.$, the number of transmit antenna in the secrecy channel estimation scheme.

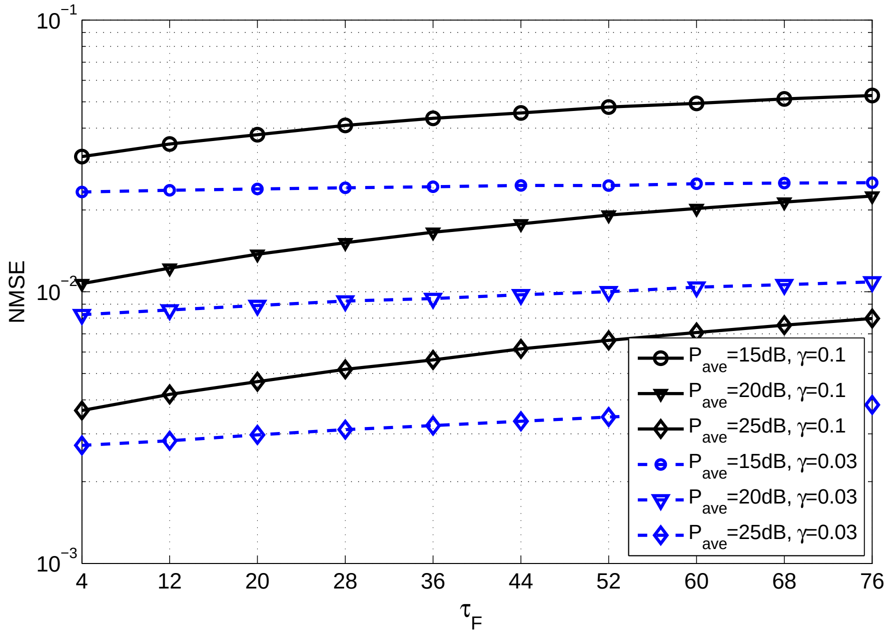


Figure 7.1: NMSE performance versus forward training τ_F for $\bar{P}_t = 30$ dB and $\bar{P}_L = 20$ dB.

and

$$\text{NMSE}_{\text{LB},nonrec} = \left(\frac{1}{\sigma_H^2} + \frac{\min\{2\bar{P}_t N_t, \; P_{ave}(3N_t + N_L)\}}{N_t \sigma_w^2} \right)^{-1} \tag{7.2}$$

which both stand for the minimum achievable NMSE at the LR when $\sigma_a^2 = 0$, $i.e.$, no AN exists.

Figure 7.1 shows the NMSE performance of LR versus the forward training length τ_F under the constant energy constraints. In the reciprocal case, the average energy constraint is given by $P_{ave}(N_t + N_L)$ and the individual energy constraints of the transmitter and LR are respectively given by $\bar{P}_t N_t$ and $\bar{P}_L N_L$. We compare different average power constraints $P_{ave} = 15$ dB, 20 dB and 25 dB and different lower limit values $\gamma = 0.1$ and 0.03. We see from Fig. 7.1 that the NMSE value of LR is monotonically non-decreasing with respect to the training length τ_F. It shows that in secrecy channel estimation it is better to keep the training length as small as possible. This is due to the fact that as the forward training length increases, it takes more AN energy to satisfy the lower limit constraint on the UR

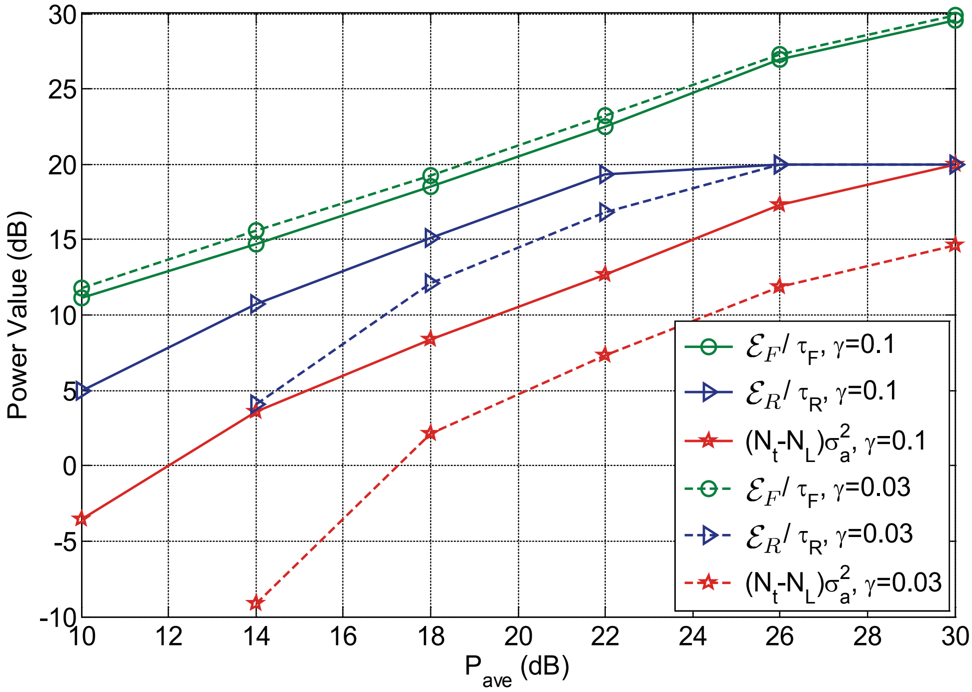


Figure 7.2: Power allocation among reverse and forward training powers $\mathcal{E}_R/\tau_R$, $\mathcal{E}_F/\tau_F$ and AN power $(N_t - N_L)\sigma_a^2$.

thus the budget for the training energies is sacrificed. Moreover, we see that the lines of $\gamma = 0.1$ are more steeper than those of $\gamma = 0.03$. The trade-off between the AN energy and training energies is more explicit as the lower limit is severer.

Figure 7.2 shows the optimal allocation of the reciprocal case among the reverse and forward training powers $\mathcal{E}_R/\tau_R$, $\mathcal{E}_F/\tau_F$ and the AN power $(N_t - N_L)\sigma_a^2$ versus average power constraint P_{ave}. We compare two different lower limit values $\gamma = 0.1$ and $\gamma = 0.03$. We see from Fig. 7.2 that it is desirable to allocate more power to the AN and less power to the forward training as γ increase from 0.03 to 0.1. This is due to the fact that the forward training signal benefits the LR and the UR equally while the AN primarily degrades the UR's estimation performance. In addition, we see that the reverse training power increases with γ, since the reverse training power mainly determines the subspace into which the AN is transmitted, which helps to reduce the interference caused by the AN on the LR. Note

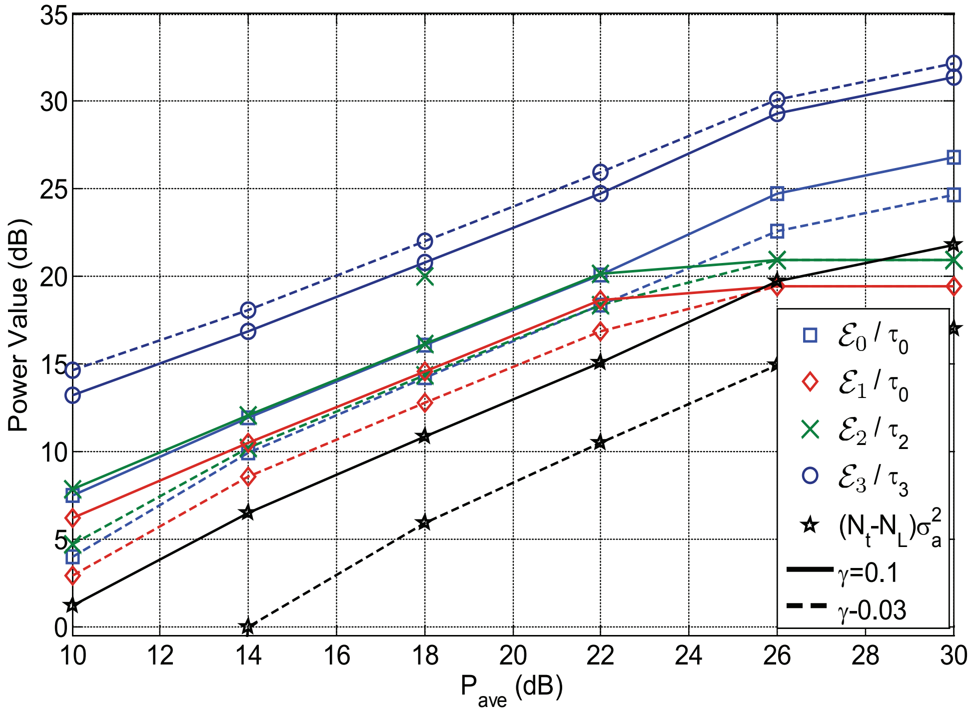


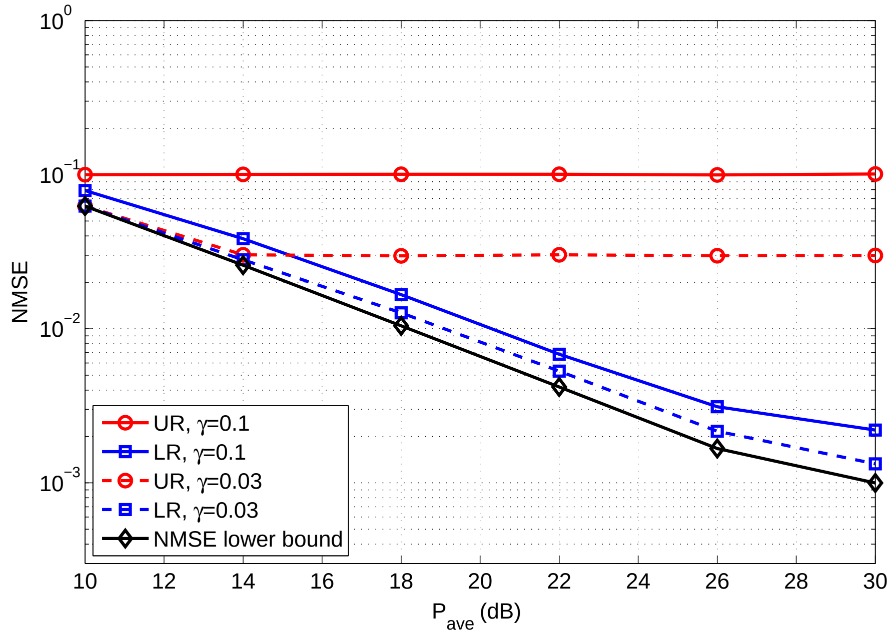
Figure 7.3: Power allocation among the training powers $\mathcal{E}_0/N_t$, $\mathcal{E}_1/N_t$ $\mathcal{E}_2/N_L$ $\mathcal{E}_3/N_t$ and AN power $(N_t - N_L)\sigma_a^2$.

that when $P_{ave} = 10$ dB and $\gamma = 0.03$, this is the case with γ out of the interesting interval (4.10), the reverse training power and AN power both equal to 0 which can not be showed in the log-value.

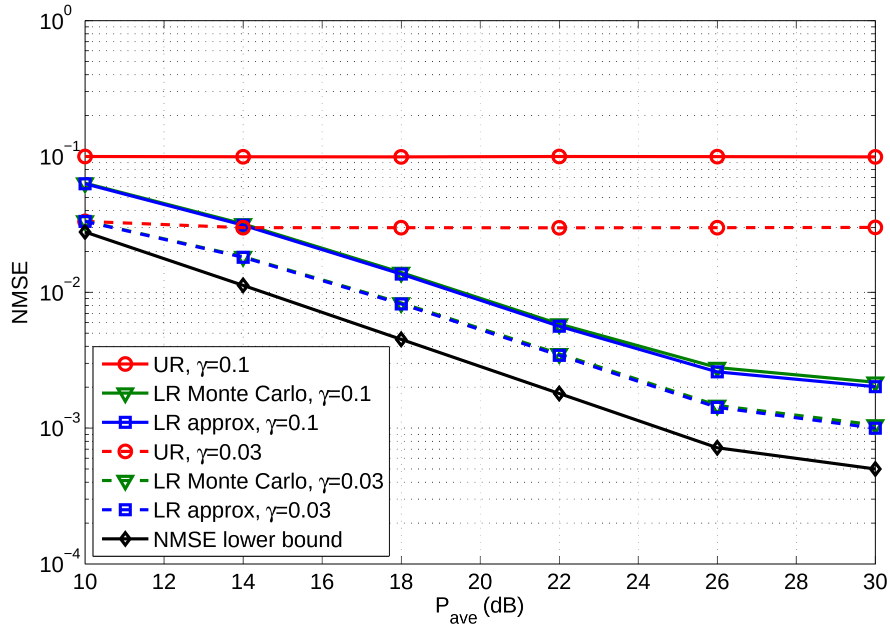
On the other hand, Figure 7.3 shows the power allocation of the non-reciprocal case among the round-trip training powers $\mathcal{E}_0/N_t$ and $\mathcal{E}_1/N_t$, reverse and forward training powers $\mathcal{E}_2/N_L$ and $\mathcal{E}_3/N_L$ and the AN power $(N_t - N_L)\sigma_a^2$ versus average power constraint P_{ave}. We have similar observation about the allocation between the forward training power and AN power to that of the reciprocal case. We see from Fig. 7.3 that the round-trip and reverse training powers all increase with respect to γ, since these powers play the role to design the placement of AN for minimizing the interference on the LR. As γ increases, so does the AN power, there needs more round-trip and reverse training powers to decrease the damage cause by the AN on the LR.

Figure 7.4(a) and Figure 7.4(b) show the NMSE performance of the LR and UR versus average power constraint P_{ave} respectively for the reciprocal and the non-reciprocal channel. We compare two different lower limit values $\gamma = 0.1$ and $\gamma = 0.03$ in both figures. From Fig. 7.4, we observed that the NMSE of the UR meets the lower limit in both reciprocal and non-reciprocal case. Furthermore, the proposed DCE scheme constrains the UR's NMSE well above γ. In addition, from Fig. 7.4(b), we see that the approximation of LR's NMSE (6.28) is quite close to the Monte-Carlo simulation result of LR's NMSE.

In Figure 7.5, we show the symbol error rate (SER) at LR and UR versus the average power constraint P_{ave} in the data transmission phase. We consider the scenario where the transmitter sends a 4×4 complex orthogonal STBC (OSTBC) with $N_t = 4$. The code length is equal to four and each code block contains three QAM source symbols [9]. The data transmission power is set to P_{ave}. Both LR and the UR will exploit their channel estimates obtained by the proposed DCE to decode the received symbols. In this Monte-Carlo simulation, the SER is obtained by averaging over 50000 channel realization and OSTBCs. In particular, Figure 7.5(a) presents the associated average SERs for 64-QAM OSTBC in the reciprocal case. We see that the SER of the LR will gradually improve while the SER of the UR remains larger than 0.1 due to the poor channel estimation performance at the UR. Figure 7.5(b) shows the associated average SERs for 64-QAM OSTBC in the non-reciprocal case. We have similar observation in this case. Both figures illustrate that, with the proposed two-way training DCE scheme, the discrimination of the data detection performances between LR and UR can be effectively achieved. It is worthwhile to mention that the feedback-and-retraining DCE scheme proposed in [3] assumes a perfect feedback channel with no power consumption and, thus, it is difficult to have a fair performance comparison between the proposed scheme and that in [3].

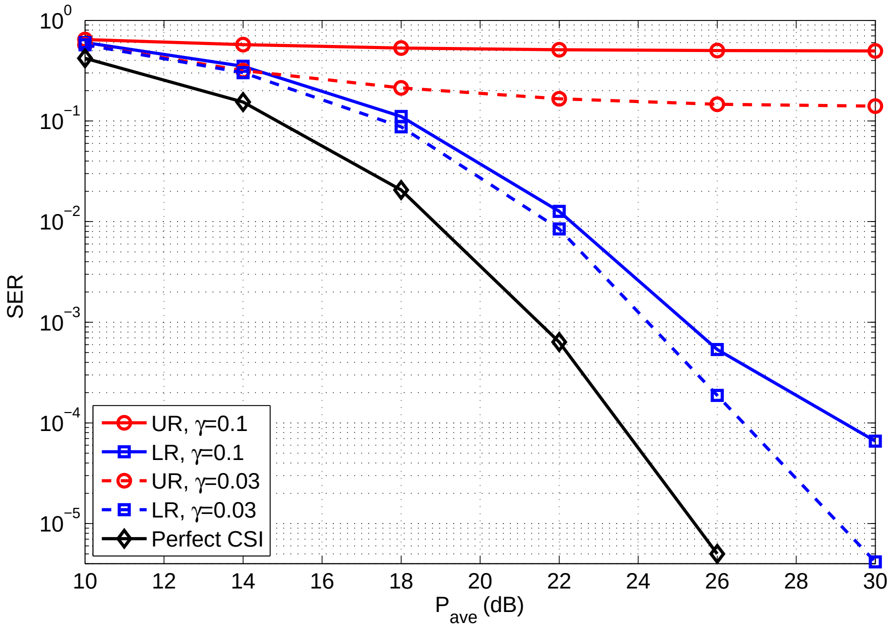


(a) Reciprocal case

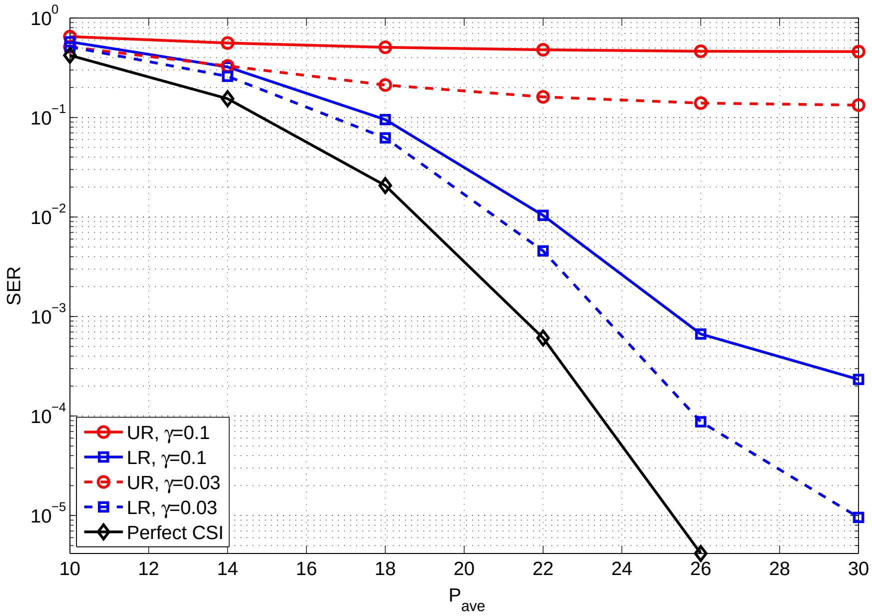


(b) Non-reciprocal case

Figure 7.4: NMSE performance of the proposed DCE scheme for the reciprocal and non-reciprocal case.



(a) Reciprocal case: 64-QAM OSTBC



(b) Non-reciprocal case: 64-QAM OSTBC

Figure 7.5: SER performance of the LR and UR in an OSTBC system with the channel estimates obtained by the proposed DCE scheme.

Chapter 8

Conclusion

In this thesis, we proposed a new DCE scheme based on the two-way training methodology, where both the transmitter and LR are allowed to emit training signals. In particular, training signals sent by LR are used to help the transmitter obtain an accurate estimate of the transmitter-to-LR channel. The proposed two-way DCE scheme utilizes two phases of training in reciprocal channels and three phases of training in non-reciprocal channels. The proposed training design drastically decreases the overall training overhead compared to the original DCE scheme proposed in [3]. The training and AN powers were optimized by minimizing the NMSE of LR subject to a preset lower limit on the NMSE of UR, an average total power constraint, and individual power constraints over all transmitters. For the case with reciprocal channels, the optimal power allocation problem was reformulated into a one-variable optimization problem which can be easily solved by simple line search. For the case with non-reciprocal channels, we derived an approximation of LR's NMSE and utilized monomial approximation and condensation method to obtain an approximate solution for the power allocation problem. Numerical results were provided to verify the efficiency of the proposed two-way DCE schemes.

Chapter 9

Appendix

9.1 Proof of Proposition I

For notational simplicity, let us define $\alpha = (N_t - N_L)\sigma_a^2$. In the following, we solve the optimization problem in two steps: (i) find the optimal values of $\mathcal{E}_F$ and α for any given $\mathcal{E}_R$; and (ii) find the optimal value of $\mathcal{E}_R$.

Step i :

Suppose a feasible $\mathcal{E}_R$ is given, the optimal values of $\mathcal{E}_F$ and α can be found as functions of $\mathcal{E}_R$ from the below optimization problem

$$\max_{\mathcal{E}_F, \alpha \geq 0} \quad \frac{(N_L\sigma_w^2 + \sigma_H^2\mathcal{E}_R)\mathcal{E}_F}{N_L\sigma_w^2 + \sigma_H^2 \cdot \mathcal{E}_R + N_L\sigma_H^2\frac{\sigma_{\tilde{w}}^2}{\sigma_w^2} \cdot \alpha} \tag{9.1a}$$

$$\text{s.t.} \quad \frac{\sigma_v^2 \cdot \mathcal{E}_F}{\sigma_G^2 \cdot \alpha + \sigma_v^2} \leq \tilde{\gamma}, \tag{9.1b}$$

$$\mathcal{E}_F + \alpha \cdot \tau_F \leq P_{ave}(\tau_R + \tau_F) - \mathcal{E}_R, \tag{9.1c}$$

$$\mathcal{E}_F + \alpha \cdot \tau_F \leq \bar{P}_t\tau_F. \tag{9.1d}$$

Note that from (4.8) a feasible $\mathcal{E}_R$ must satisfy $\mathcal{E}_R \leq \bar{P}_L\tau_R$. In the following, we consider two different ranges of $\mathcal{E}_R$.

Case 1 ($P_{ave}(\tau_R + \tau_F) - \tilde{\gamma} < \mathcal{E}_R \leq \bar{P}_L\tau_R$): If $P_{ave}(\tau_R + \tau_F) - \tilde{\gamma} < \bar{P}_L\tau_R$ holds. Since the

objective function in (9.1) is monotonically increasing with respect to $\mathcal{E}_F$ but decreasing with respect to α, by (4.12) and the condition of $\mathcal{E}_R > P_{ave}(\tau_R + \tau_F) - \tilde{\gamma}$, we get $\mathcal{E}_F^*(\mathcal{E}_R) = P_{ave}(\tau_R + \tau_F) - \mathcal{E}_R$, $\alpha^* = 0$ and hence the value of (9.1a) becomes

$$\mathcal{E}_F^*(\mathcal{E}_R) = P_{ave}(\tau_R + \tau_F) - \mathcal{E}_R, \tag{9.2}$$

which is less than $\tilde{\gamma}$.

Case 2 ($\mathcal{E}_R \leq \min\{\bar{P}_L T_R,\ P_{ave}(\tau_R + \tau_F) - \tilde{\gamma}\}$): It can be observed that if the constraint (9.1b) is inactive we can always decrease α until activating the constraint to obtain a larger objective value. If the condition (9.1b) is still inactive even when $\alpha = 0$, we can instead lift $\mathcal{E}_F$ to achieve a larger objective value while still satisfying (4.9), (4.12) and the condition $\mathcal{E}_R \leq \min\{\bar{P}_L T_R,\ P_{ave}(\tau_R + \tau_F) - \tilde{\gamma}\}$. We conclude that constraint (9.1b) must be active at the optimum. Hence we have

$$\mathcal{E}_F^*(\mathcal{E}_R) = \tilde{\gamma}\left(\frac{\sigma_G^2}{\sigma_v^2} \cdot \alpha^*(\mathcal{E}_R) + 1\right). \tag{9.3}$$

By substituting (9.3) into (9.1), the problem becomes

$$\begin{aligned}
\max_{\alpha \geq 0} \quad & \frac{(\sigma_G^2/\sigma_v^2 \cdot \alpha + 1)(N_L \sigma_{\tilde{w}}^2 + \sigma_H^2 \cdot \mathcal{E}_R)\tilde{\gamma}}{N_L \sigma_H^2 \frac{\sigma_{\tilde{w}}^2}{\sigma_w^2} \cdot \alpha + N_L \sigma_{\tilde{w}}^2 + \sigma_H^2 \cdot \mathcal{E}_R} && \text{(9.4a)}\\
\text{s.t.} \quad & \left(\tau_F + \frac{\sigma_G^2 \tilde{\gamma}}{\sigma_v^2}\right)\alpha + \mathcal{E}_R \leq P_{ave}(\tau_R + \tau_F) - \tilde{\gamma} && \text{(9.4b)}\\
& \tilde{\gamma}\left(\frac{\sigma_G^2}{\sigma_v^2}\alpha + 1\right) + \tau_F \cdot \alpha \leq \bar{P}_t \tau_F. && \text{(9.4c)}
\end{aligned}$$

The range of $\mathcal{E}_R$ in this case is further divided into the following two subranges.

(a) When $\mathcal{E}_R < \mu$ and $\mathcal{E}_R \leq \min\{\bar{P}_L \tau_R,\ P_{ave}(\tau_R + \tau_F) - \tilde{\gamma}\}$, where $\mu \triangleq N_L\left(\frac{\sigma_v^2 \sigma_{\tilde{w}}^2}{\sigma_G^2 \sigma_w^2} - \frac{\sigma_{\tilde{w}}^2}{\sigma_H^2}\right)$, the objective function in (9.4a) is a monotonically decreasing function with respect to α. Therefore, the optimal value of $\alpha^*(\mathcal{E}_R)$ is 0 and the corresponding optimal objective value is equal to $\tilde{\gamma}$.

(b) When $\mu \leq \mathcal{E}_R \leq \min\{\bar{P}_L\tau_R,\ P_{ave}(\tau_R+\tau_F)-\tilde{\gamma}\}$, the objective function in (9.4a) is monotonically non-decreasing with respect to α. For $\mu \leq \mathcal{E}_R \leq P_{ave}(\tau_R+\tau_F)-\bar{P}_t\tau_F$, constraint (9.4c) must be active at the optimum with

$$\alpha^* = \frac{\bar{P}_t\tau_F - \tilde{\gamma}}{\tau_F + \sigma_G^2\tilde{\gamma}/\sigma_v^2}. \tag{9.5}$$

Reversely, considering $\mathcal{E}_R \geq \max\{\mu,\ P_{ave}(\tau_R+\tau_F)-\bar{P}_t\tau_F\}$, the constraint (9.4b) must be active at the optimum with

$$\alpha^*(\mathcal{E}_R) = \frac{P_{ave}(\tau_R+\tau_F) - \tilde{\gamma} - \mathcal{E}_R}{\tau_F + \sigma_G^2\tilde{\gamma}/\sigma_v^2} \tag{9.6}$$

Moreover, for $\mathcal{E}_R \geq \mu$, the optimal objective value of (9.4a) can be shown to be

$$\frac{\sigma_G^2 c^*/\sigma_v^2 + 1}{\frac{N_L\sigma_H^2\sigma_{\tilde{w}}^2/\sigma_w^2}{N_L\sigma_w^2+\sigma_H^2\cdot a}c^* + 1}\cdot\tilde{\gamma} \geq \tilde{\gamma}. \tag{9.7}$$

which is no less than $\tilde{\gamma}$.

Step ii :

We now solve for the optimal value of $\mathcal{E}_R$. From the analysis in Step i, a feasible $\mathcal{E}_R$ satisfying $\mathcal{E}_R \leq \min\{\bar{P}_L\tau_R,\ P_{ave}(\tau_R+\tau_F)-\tilde{\gamma}\}$ leads to greater objective value than that of $P_{ave}(\tau_R+\tau_F)-\tilde{\gamma} < \mathcal{E}_R \leq \bar{P}_L\tau_R$, thus the optimal value of $\mathcal{E}_R$ must lie in the former condition. For the first case that $\mu > \min\{\bar{P}_L\tau_R,\ P_{ave}(\tau_R+\tau_F)-\tilde{\gamma}\}$, we can infer $\mathcal{E}_R < \mu$ for all feasible $\mathcal{E}_R$ satisfying $\mathcal{E}_R \leq \min\{\bar{P}_L\tau_R,\ P_{ave}(\tau_R+\tau_F)-\tilde{\gamma}\}$ so that $\alpha^* = 0$ and thus $\mathcal{E}_F^* = \tilde{\gamma}$. Then we get $\mathcal{E}_R^* = 0$ for no need of AN. For the other case of $\mu \leq \min\{\bar{P}_L\tau_R,\ P_{ave}(\tau_R+\tau_F)-\tilde{\gamma}\}$, from (9.2) and (9.7) we can see that the corresponding objective value for $\max\{0,\ \mu\} \leq \mathcal{E}_R \leq \min\{\bar{P}_L\tau_R,\ P_{ave}(\tau_R+\tau_F)-\tilde{\gamma}\}$ is no less than that for $\mathcal{E}_R < \mu$. If $\mu \leq \mathcal{E}_R \leq P_{ave}(\tau_R+\tau_F)-\bar{P}_t\tau_F$ exists, the optimization problem (9.4) becomes

$$\begin{aligned} \max_{\widetilde{\mathcal{E}}_R \geq 0} \quad & \frac{(N_L\sigma_{\tilde{w}}^2 + \sigma_H^2\widetilde{\mathcal{E}}_R)\mathcal{E}_F^\star}{N_L\sigma_{\tilde{w}}^2 + \sigma_H^2\cdot\widetilde{\mathcal{E}}_R + N_L\sigma_H^2\frac{\sigma_{\tilde{w}}^2}{\sigma_w^2}\cdot\alpha^\star} \\ \text{s.t} \quad & \max\{0,\ \mu\} \leq \widetilde{\mathcal{E}}_R \leq P_{ave}(\tau_R+\tau_F) - \bar{P}_t\tau_F \end{aligned} \tag{9.8}$$

where $\mathcal{E}_F^\star$ and $\alpha^\star$ are given by (9.3) and (9.5) which do not depend on $\mathcal{E}_R$ in this condition. It can be observed that the objective function (9.8) is monotonically non-decreasing with respect to $\tilde{a}$; thus the optimal value is achieved when $\widetilde{\mathcal{E}}_R^* = P_{ave}(\tau_R + \tau_F) - \bar{P}_t\tau_F$. However, the corresponding optimal objective value of (9.8) is the same as the objective value of (4.13) in this case. Hence, we can have the value of $\mathcal{E}_R^*$ lie in the interval $\max\{0,\ \mu,\ P_{ave}(\tau_R+\tau_F) - \bar{P}_t\tau_F\} \le \mathcal{E}_R \le \min\{\bar{P}_L\tau_R,\ P_{ave}(\tau_R+\tau_F) - \tilde{\gamma}\}$ by solving the optimization problem (4.13) and the corresponding $\alpha(\mathcal{E}_R)$ and $\mathcal{E}_F(\mathcal{E}_R)$ are given by (9.6) and (9.3), respectively.

9.2 Monomial Approximation and Condensation Method for the Problem in (6.33)

Here, we show how to obtain an efficient solution for the problem in (6.33) using monomial approximation and condensation method. The problem in (6.33) can be stated as follows:

$$\min_{\mathcal{E}_0,\mathcal{E}_1,\mathcal{E}_2,\mathcal{E}_3,\sigma_a^2\ge 0} \quad \left(\frac{1}{\sigma_{H_d}^2} + \frac{1}{\sigma_w^2}\frac{f_1(\mathcal{E}_0,\alpha^2,\mathcal{E}_2,\mathcal{E}_3)}{f_2(\mathcal{E}_0,\alpha^2,\mathcal{E}_2,\mathcal{E}_3)}\right)^{-1} \tag{9.9a}$$

$$\text{subject to} \quad \frac{\mathcal{E}_3}{N_t}\frac{1}{\sigma_G^2(N_t-N_L)\sigma_a^2+\sigma_w^2} \le \frac{1}{\gamma} - \frac{1}{\sigma_G^2} \tag{9.9b}$$

$$\mathcal{E}_0 + \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 + (N_t - N_L)\sigma_a^2 N_t \le P_{ave}(N_t + N_t + N_L + N_t) \tag{9.9c}$$

$$\mathcal{E}_0 + \mathcal{E}_3 + (N_t - N_L)\sigma_a^2 N_t \le \bar{P}_t(N_t + N_t) \tag{9.9d}$$

$$\mathcal{E}_1 + \mathcal{E}_2 \le \bar{P}_L(N_t + N_L) \tag{9.9e}$$

where

$$f_1(\mathcal{E}_0,\alpha^2,\mathcal{E}_2,\mathcal{E}_3) = \frac{\mathcal{E}_3}{N_t}\left(N_L\left(\frac{\sigma_{H_d}^2\mathcal{E}_0}{N_t}+\sigma_w^2\right)\alpha^2 + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}\left(\frac{\sigma_{H_d}^2\mathcal{E}_0}{N_t}+\sigma_w^2\right)\alpha^2\frac{\mathcal{E}_2}{N_L} + \frac{\mathcal{E}_2}{N_L} + \frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2}\right)$$

and

$$
\begin{aligned}
&f_2(\mathcal{E}_0,\alpha^2,\mathcal{E}_2,\mathcal{E}_3)\\
&\quad=(N_t-N_L)\sigma_a^2\sigma_{H_d}^2\left(\frac{N_L}{\sigma_w^2}\left(\frac{\sigma_{H_d}^2\mathcal{E}_0}{N_t}+\sigma_w^2\right)\alpha^2+\frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}\alpha^2\frac{\mathcal{E}_2}{N_L}+\frac{\mathcal{E}_2}{N_L\sigma_w^2}+\frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2\sigma_w^2}\right)\\
&\qquad+\left(N_L\left(\frac{\sigma_{H_d}^2\mathcal{E}_0}{N_t}+\sigma_w^2\right)\alpha^2+\frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}\left(\frac{\sigma_{H_d}^2\mathcal{E}_0}{N_t}+\sigma_w^2\right)\alpha^2\frac{\mathcal{E}_2}{N_L}+\frac{\mathcal{E}_2}{N_L}+\frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2}\right).
\end{aligned}
$$

By introducing the auxiliary variable

$$
t=\frac{f_1(\mathcal{E}_0,\alpha^2,\mathcal{E}_2,\mathcal{E}_3)}{f_2(\mathcal{E}_0,\alpha^2,\mathcal{E}_2,\mathcal{E}_3)} \tag{9.10}
$$

and by defining the variables $t_0=\frac{\sigma_{H_d}^2\mathcal{E}_0}{N_t}+\sigma_w^2$, $t_1=\alpha^2$, $t_2=\frac{\mathcal{E}_2}{N_L}$, $t_3=\frac{\mathcal{E}_3}{N_t}$, and $t_4=(N_t-N_L)\sigma_a^2\sigma_G^2+\sigma_v^2$, the problem can be reformulated as

$$
\begin{aligned}
&\min_{t,t_0,t_1,t_2,t_3,t_4\geq 0} && t^{-1} && \text{(9.11a)}\\
&\text{subject to} && t\leq\frac{\bar{f}_1(t_0,t_1,t_2,t_3)}{\bar{f}_2(t_0,t_1,t_2,t_3)} && \text{(9.11b)}\\
& && \sigma_w^2t_0^{-1}\leq 1 && \text{(9.11c)}\\
& && \sigma_v^2t_4^{-1}\leq 1 && \text{(9.11d)}\\
& && c_1t_3t_4^{-1}\leq 1 && \text{(9.11e)}\\
& && c_2\left(\frac{N_t}{\sigma_{H_d}^2}t_0+N_tN_Lt_0t_1+N_Lt_2+N_tt_3+\frac{N_t}{\sigma_G^2}t_4\right)\leq 1 && \text{(9.11f)}\\
& && c_3\left(\frac{N_t}{\sigma_{H_d}^2}t_0+N_tt_3+\frac{N_t}{\sigma_G^2}t_4\right)\leq 1 && \text{(9.11g)}\\
& && c_4(N_tN_Lt_0t_1+N_Lt_2)\leq 1 && \text{(9.11h)}
\end{aligned}
$$

where

$$
\bar{f}_1(t_0,t_1,t_2,t_3)=N_Lt_0t_1t_3+\frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}t_0t_1t_2t_3+t_2t_3+\frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2}t_3,
$$

$$
\begin{aligned}
\bar{f}_2(t_0,t_1,t_2,t_3)=&\left(\frac{t_4}{\sigma_G^2}-\frac{\sigma_v^2}{\sigma_G^2}\right)\sigma_{H_d}^2\left(\frac{N_L}{\sigma_w^2}t_0t_1+\frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}t_1t_2+\frac{1}{\sigma_w^2}t_2+\frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2\sigma_w^2}\right)\\
&+N_Lt_0t_1+\frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}t_0t_1t_2+t_2+\frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2}
\end{aligned}
$$

and

$$c_1 = \left(\frac{1}{\gamma} - \frac{1}{\sigma_G^2}\right)^{-1}, \quad c_2 = \left(P_{ave}(3N_t + N_L) + \frac{N_t\sigma_w^2}{\sigma_{H_d}^2} + \frac{\sigma_v^2 N_t}{\sigma_G^2}\right)$$
$$c_3 = \left(2\bar{P}_t N_t + \frac{N_t\sigma_w^2}{\sigma_{H_d}^2} + \frac{\sigma_v^2 N_t}{\sigma_G^2}\right), \quad c_4 = \left(\bar{P}_L(N_t + N_L)\right)^{-1}.$$

Note that in (9.11b) the equality was replaced by the inequality since one can inspect that the inequality must be active when the optimal objective value is achieved. To make sure $\mathcal{E}_0$ and σ_a^2 are no less than zero, we attach two posynomial constraints (9.11c) and (9.11d). In addition, we have reformulated the constraints in (9.11e-9.11h) into posynomial inequalities, which are standard inequality constraints for GP. However, the inequality constraint in (9.11b) is not a standard GP inequality. It can only be expressed as a ratio of posynomials as given below:

$$\frac{\frac{\sigma_{H_d}^2}{\sigma_G^2}\left(\frac{N_L}{\sigma_w^2}t_0t_1t_4t + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}t_1t_2t_4t + \frac{1}{\sigma_w^2}t_2t_4t + \frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2\sigma_w^2}t_4t\right) + N_Lt_0t_1t + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}t_0t_1t_2t + t_2t + \frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2}t}{\frac{\sigma_v^2\sigma_{H_d}^2}{\sigma_G^2}\left(\frac{N_L}{\sigma_w^2}t_0t_1 + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}t_1t_2 + \frac{1}{\sigma_w^2}t_2 + \frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2\sigma_w^2}\right) + N_Lt_0t_1t_3 + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}t_0t_1t_2t_3 + t_2t_3 + \frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2}t_3} \le 1 \tag{9.12}$$

In order to simplify the problem into a standard GP form, we apply the monomial approximation [14] to transform this into a posynomial constraint. In particular, if a set of feasible points $\{\bar{t}, \bar{t}_1, \bar{t}_2, \bar{t}_3, \bar{t}_4\}$ of problem (9.11) is given, the inequality in (9.12) can be replaced by the posynomial constraint given below:

$$\frac{\frac{\sigma_{H_d}^2}{\sigma_G^2}\left(\frac{N_L}{\sigma_w^2}t_0t_1t_4t + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}t_1t_2t_4t + \frac{1}{\sigma_w^2}t_2t_4t + \frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2\sigma_w^2}t_4t\right) + N_Lt_0t_1t + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}t_0t_1t_2t + t_2t + \frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2}t}{g(\bar{t}, \bar{t}_1, \bar{t}_2, \bar{t}_3)\left(\frac{t_0}{\bar{t}_0}\right)^{\theta_0}\left(\frac{t_1}{\bar{t}_1}\right)^{\theta_1}\left(\frac{t_2}{\bar{t}_2}\right)^{\theta_2}\left(\frac{t_3}{\bar{t}_3}\right)^{\theta_3}} \le 1 \tag{9.13}$$

where

$$g(\bar{t}, \bar{t}_1, \bar{t}_2, \bar{t}_3) = \frac{\sigma_v^2\sigma_{H_d}^2}{\sigma_G^2}\left(\frac{N_L}{\sigma_w^2}t_0t_1 + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}t_1t_2 + \frac{1}{\sigma_w^2}t_2 + \frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2\sigma_w^2}\right)$$
$$+ N_Lt_0t_1t_3 + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}t_0t_1t_2t_3 + t_2t_3 + \frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2}t_3,$$

$$\theta_0 = \frac{\frac{\sigma_v^2\sigma_{H_d}^2 N_L}{\sigma_G^2\sigma_{\tilde{w}}^2}\bar{t}_0\bar{t}_1 + N_L\bar{t}_0\bar{t}_1\bar{t}_3 + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}\bar{t}_0\bar{t}_1\bar{t}_2\bar{t}_3}{g(\bar{t},\bar{t}_1,\bar{t}_2,\bar{t}_3)},$$

$$\theta_1 = \frac{\frac{\sigma_v^2\sigma_{H_d}^2 N_L}{\sigma_G^2\sigma_{\tilde{w}}^2}\bar{t}_0\bar{t}_1 + \frac{\sigma_v^2\sigma_{H_d}^2\sigma_{H_u}^2 N_t}{\sigma_G^2\sigma_{\tilde{w}}^2}\bar{t}_1\bar{t}_2 + N_L\bar{t}_0\bar{t}_1\bar{t}_3 + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}\bar{t}_0\bar{t}_1\bar{t}_2\bar{t}_3}{g(\bar{t},\bar{t}_1,\bar{t}_2,\bar{t}_3)},$$

$$\theta_2 = \frac{\frac{\sigma_v^2\sigma_{H_d}^2\sigma_{H_u}^2 N_t}{\sigma_G^2\sigma_{\tilde{w}}^2}\bar{t}_1\bar{t}_2 + \frac{\sigma_v^2\sigma_{H_d}^2}{\sigma_G^2\sigma_{\tilde{w}}^2}\bar{t}_2 + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}\bar{t}_0\bar{t}_1\bar{t}_2\bar{t}_3 + \bar{t}_2\bar{t}_3}{g(\bar{t},\bar{t}_1,\bar{t}_2,\bar{t}_3)},$$

and

$$\theta_3 = \frac{N_L\bar{t}_0\bar{t}_1\bar{t}_3 + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}\bar{t}_0\bar{t}_1\bar{t}_2\bar{t}_3 + \bar{t}_2\bar{t}_3 + \frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2}\bar{t}_3}{g(\bar{t},\bar{t}_1,\bar{t}_2,\bar{t}_3)}$$

Hence, for a given set of feasible points $\{\bar{t},\bar{t}_1,\bar{t}_2,\bar{t}_3,\bar{t}_4\}$, the problem in (9.11) can be approximate by the following problem

$$\min_{e_0,e_1,e_2,e_3,\sigma_a^2\geq 0} \quad t^{-1} \tag{9.14a}$$

$$\text{subject to} \quad \frac{\frac{\sigma_{H_d}^2}{\sigma_G^2}\left(\frac{N_L}{\sigma_{\tilde{w}}^2}t_0t_1t_4t + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}t_1t_2t_4t + \frac{1}{\sigma_{\tilde{w}}^2}t_2t_4t + \frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2\sigma_{\tilde{w}}^2}t_4t\right)}{g(\bar{t},\bar{t}_1,\bar{t}_2,\bar{t}_3)\left(\frac{t_0}{\bar{t}_0}\right)^{\theta_0}\left(\frac{t_1}{\bar{t}_1}\right)^{\theta_1}\left(\frac{t_2}{\bar{t}_2}\right)^{\theta_2}\left(\frac{t_3}{\bar{t}_3}\right)^{\theta_3}} + \frac{N_Lt_0t_1t + \frac{N_t\sigma_{H_u}^2}{\sigma_{\tilde{w}}^2}t_0t_1t_2t + t_2t + \frac{\sigma_{\tilde{w}}^2}{\sigma_{H_u}^2}t}{g(\bar{t},\bar{t}_1,\bar{t}_2,\bar{t}_3)\left(\frac{t_0}{\bar{t}_0}\right)^{\theta_0}\left(\frac{t_1}{\bar{t}_1}\right)^{\theta_1}\left(\frac{t_2}{\bar{t}_2}\right)^{\theta_2}\left(\frac{t_3}{\bar{t}_3}\right)^{\theta_3}} \leq 1 \tag{9.14b}$$

$$\sigma_w^2t_0^{-1} \leq 1 \tag{9.14c}$$

$$\sigma_v^2t_4^{-1} \leq 1 \tag{9.14d}$$

$$c_1t_3t_4^{-1} \leq 1 \tag{9.14e}$$

$$c_2\left(\frac{N_t}{\sigma_{H_d}^2}t_0 + N_tN_Lt_0t_1 + N_Lt_2 + N_tt_3 + \frac{N_t}{\sigma_G^2}t_4\right) \leq 1 \tag{9.14f}$$

$$c_3\left(\frac{N_t}{\sigma_{H_d}^2}t_0 + N_tt_3 + \frac{N_t}{\sigma_G^2}t_4\right) \leq 1 \tag{9.14g}$$

$$c_4(N_tN_Lt_0t_1 + N_Lt_2) \leq 1 \tag{9.14h}$$

The problem then becomes a standard GP problem and can be efficiently solved by a few simple computer softwares such as `CVX` [15]. The condensation method then proposes to

repeat this process iteratively by replacing the set of feasible points $\{\bar{t}, \bar{t}_1, \bar{t}_2, \bar{t}_3, \bar{t}_4\}$ in each iteration with the optimal solution of (9.14) obtained in the previous iteration. This process continues until no further improvements can be obtained.

Bibliography

[1] P. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 10, pp. 4687-4698, Oct. 2008.

[2] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas: the MISOME wiretap channel", *IEEE Trans. Inform. Theory,* vol. 56, no. 7, pp. 3088-3104, Jul. 2010.

[3] T.-H. Chang, W.-C. Chiang, Y.-W. Hong and C.-Y. Chi, "Training sequence design for discriminatory channel estimation in wireless MIMO systems," *IEEE Trans. Signal Process.,* vol. 58, no. 12, pp. 6223-6237, Dec. 2010.

[4] C. Steger and A. Sabharwal, "Single-input two-way SIMO Channel: diversity-multiplexing tradeoff with two-way training," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4877-4885, Dec. 2008.

[5] X. Zhou, T. A. Lamahewa, P. Sadeghi and S. Durrani, "Two-way training: optimal power allocatin for pilot and data transmission," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 564-569, Feb. 2010.

[6] L. P. Withers, R. M. Taylor and D. M. Warme, "Echo-MIMO: a two-way channel training method for matched cooperative beamforming," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4419-4432, Sep. 2008.

[7] S. M. Kay, *Fundamnetals of Statistical Signal Processing: Estimation Theory.* New Jersey: Prentice Hall International, 1993.

[8] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE. Trans. Inform. Theory*, vol. 49, no. 4, pp. 951-963, Apr. 2003.

[9] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications.* Cambridge, UK: Cambridge University Press, 2003.

[10] A. Gupta and D. Nagar, *Matrix variate distributions.* Chapman & Hall/CRC, 2000.

[11] A. Edelman and N. Rao, "Random matrix theory," *Acta Numerica*, vol. 14, pp. 233-297, 2005.

[12] A. Zanella, M. Chiani and M. Z. Win, "On the marginal distribution of the eigenvalues of Wishart matrices," *IEEE Trans. Wireless Commun.*, vol. 57, no. 4, pp. 1050-1060, April 2009.

[13] C. Martin and B. Ottersten, "Asymptotic eigenvalue distributions and capacity for MIMO channels under correlated fading," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1350-1359, July 2004.

[14] S. Boyd, S.-J. Kim, L. Vandenberghe and A. Hassibi, "A tutorial on geometric programming," *Optim. Eng.*, vol. 8, pp. 67-127, April 2007.

[15] M. Grant and S. Boyd. "CVX: Matlab software for disciplined convex programming," version 1.21. http://cvxr.com/cvx, April 2011.